

	Matches	738;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	ATACGACTCACTATAGCGCGAATTCGAGCTGGGTACCGGGGATCCTCTAGAGTCGACCT	60							
Db	1	ATACGACTCACTATAGCGCGAATTCGAGCTGGGTACCGGGGATCCTCTAGAGTCGACCT	60							
QY	61	GCAGGCGATGCAAGCTTGAGTATCTATAGTTCACCTAAATAGCTTGGCGTATCATGGT	120							
Db	61	GCAGGCGATGCAAGCTTGAGTATCTATAGTTCACCTAAATAGCTTGGCGTATCATGGT	120							
QY	121	CATAGCTGTTTCTCTGTGTGAATTTGATATCCGGTCACAATTCACACAACAACGAGCGG	180							
Db	121	CATAGCTGTTTCTCTGTGTGAATTTGATATCCGGTCACAATTCACACAACAATAGAGCGG	180							
QY	181	GAAGCATAAAGTGTAAAGCTCGGGTGCCCTAATGAGTGAGCTAACTCACAATTAATTCGGT	240							
Db	181	GAAGCATAAAGTGTAAAGCTCGGGTGCCCTAATGAGTGAGCTAACTCACAATTAATTCGGT	240							
QY	241	TGCGCTCACTGCCCGCTTTCCAGTCGGGAACCTGTCGTGCCAGCTGCATTAAATGAATCG	300							
Db	241	TGCGCTCACTGCCCGCTTTCCAGTCGGGAACCTGTCGTGCCAGCTGCATTAAATGAATCG	300							
QY	301	GCCAAACGCGCGGGAGAGCGGTTTGGCTATTGGCGGCTCTTCCTCGCTCACTG	360							
Db	301	GCCAAACGCGCGGGAGAGCGGTTTGGCTATTGGCGGCTCTTCCTCGCTCACTG	360							
QY	361	ACTCGTCTCGGCTCGGTCGTTGCGGCTCGCGGAGCGGATATCAGCTCACTCAAAAGCGGTAA	420							
Db	361	ACTCGTCTCGGCTCGGTCGTTGCGGCTCGCGGAGCGGATATCAGCTCACTCAAAAGCGGTAA	420							
QY	421	TACGGTTATCCACAGATCAGGGGATTAACGCAGGAAGAACATGTGAGCAAAAGGCCAGC	480							
Db	421	TACGGTTATCCACAGATCAGGGGATTAACGCAGGAAGAACATGTGAGCAAAAGGCCAGC	480							
QY	481	AAAAGGCCAGGAACCGTAAAAAGCCGCTTGTGGCGTTTTCATAGGCTCCGCCGCC	540							
Db	481	AAAAGGCCAGGAACCGTAAAAAGCCGCTTGTGGCGTTTTCATAGGCTCCGCCGCC	540							
QY	541	CTGACGAGCATCAAAAATCGAGGCTCAAGTCAGAGGTGGCAAAACCCGACAGGACTAT	600							
Db	541	CTGACGAGCATCAAAAATCGAGGCTCAAGTCAGAGGTGGCAAAACCCGACAGGACTAT	600							
QY	601	AAAGATPACGAGCGTTTCCCGCTGGAAGCTCCCTCGTGGCGCTCTCTGTGTCCGACCTGC	660							
Db	601	AAAGATPACGAGCGTTTCCCGCTGGAAGCTCCCTCGTGGCGCTCTCTGTGTCCGACCTGC	660							
QY	661	CGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCT	720							
Db	661	CGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCT	720							
QY	721	CACGCTGTAGGTATCTCA	738							
Db	721	CACGCTGTAGGTATCTCA	738							

RESULT	2
ARI141144	
LOCUS	
DEFINITION	ARI141144
ACCESSION	Sequence 3 from patent US 6145434.
VERSION	ARI141144.1
GI:	GI:15100661
Linear	738 bp DNA PAT 08-AUG-2001

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES

Unknown.
Unknown.
Unclassified.
Unclassified.
1 (bases 1 to 738)
Tanaka, T. and Kobayashi, K.
Stencil printing method and device
Patent: US 6145334 A 3 4-NOV-2000;
Location/Qualifiers
1 738
source

BASE COUNT	170 a	208 c	192 g	168 t
/organism="unknown"				

ORIGIN	Query Match	100.0%	Score 738	DB 6	Length 738	
	Best Local Similarity	100.0%	Pred. No. 9.7e-198			
	Matches 738	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
QY	1	ATACGACTCACTATAGAGGCGAATTCGAGCTCGGTACCCGGGAGTCCCTCTAGAGTCGACCT	60			
Db	1	ATACGACTCACTATAGAGGCGAATTCGAGCTCGGTACCCGGGAGTCCCTCTAGAGTCGACCT	60			
QY	61	GCAGGCATGCAAGCTTGAGTATTCATAGTGTACCTAAATAGCTTTGGCGTAATCATGGT	120			
Db	61	GCAGGCATGCAAGCTTGAGTATTCATAGTGTACCTAAATAGCTTTGGCGTAATCATGGT	120			
QY	121	CATAGCTGTTTCCCTGCTGTGAATTTGTTATCCGCTCACAATTCACATAACATACGAGCCG	180			
Db	121	CATAGCTGTTTCCCTGCTGTGAATTTGTTATCCGCTCACAATTCACATAACATACGAGCCG	180			
QY	181	GAAGCATAAAGTGTAAAGCCTGGGGTGCTTAATGAGTGAAGTCACTCAATTAATGAATCG	240			
Db	181	GAAGCATAAAGTGTAAAGCCTGGGGTGCTTAATGAGTGAAGTCACTCAATTAATGAATCG	240			
QY	241	TGCGCTCACTGCCCGCTTTCAGTCGGGAACCTGTCGTGCGAGTCACTCAAAAGCGGTAA	300			
Db	241	TGCGCTCACTGCCCGCTTTCAGTCGGGAACCTGTCGTGCGAGTCACTCAAAAGCGGTAA	300			
QY	301	GCCAGCGCGGGGAGAGCGGTTTGCGTATTTGGGCGCTCTTCCGCTTCCCTCGCTCACTG	360			
Db	301	GCCAGCGCGGGGAGAGCGGTTTGCGTATTTGGGCGCTCTTCCGCTTCCCTCGCTCACTG	360			
QY	361	ACTCGCTGGCTCGGTCGTCGTCGGCTGGGCGAGCGGTTTGCGTATTTGGGCGCTCTTCCGCTTCCCTCGCTCACTG	420			
Db	361	ACTCGCTGGCTCGGTCGTCGTCGGCTGGGCGAGCGGTTTGCGTATTTGGGCGCTCTTCCGCTTCCCTCGCTCACTG	420			
QY	421	TACGGTTATCCACAAATCAGGGATTAACGACGAAAGACATGTGAGCAAAAGCGGTA	480			
Db	421	TACGGTTATCCACAAATCAGGGATTAACGACGAAAGACATGTGAGCAAAAGCGGTA	480			
QY	481	AAAAGGCCAGAACCGTAAAGGCGCGGTTGCTGCGGTTTTTCCATAGGCTCCGCGCCC	540			
Db	481	AAAAGGCCAGAACCGTAAAGGCGCGGTTGCTGCGGTTTTTCCATAGGCTCCGCGCCC	540			
QY	541	CTGACGAGCATCACAAAATCGACCTCAAGTCAGAGTGGCGAAACCCGACGAGGACTAT	600			
Db	541	CTGACGAGCATCACAAAATCGACCTCAAGTCAGAGTGGCGAAACCCGACGAGGACTAT	600			
QY	601	AAAGATACCGAGCGGTTTCCCGCTGGAAGTCCCTCGTGGCGTCTCCTGTTCCGACCTGC	660			
Db	601	AAAGATACCGAGCGGTTTCCCGCTGGAAGTCCCTCGTGGCGTCTCCTGTTCCGACCTGC	660			
QY	661	CGCTTACCGGATACCTGTCGCGCTTTCTCCCTCGGGAAGCGTGGCGCTTCTCATAGCT	720			
Db	661	CGCTTACCGGATACCTGTCGCGCTTTCTCCCTCGGGAAGCGTGGCGCTTCTCATAGCT	720			
QY	721	CAGCTGTAGGTATCTCA	738			
Db	721	CAGCTGTAGGTATCTCA	738			
RESULT	3					
LOCUS	AX046223					
DEFINITION	Sequence 3 from Patent WO0066771.	972 bp	DNA	linear	PAT 24-NOV-2000	
ACCESSION	AX046223					
VERSION	AX046223.1					
KEYWORDS						
SOURCE						
ORGANISM	synthetic construct.					
REFERENCE	synthetic construct.					
AUTHORS	artificial sequence.					
TITLE	1 (bases 1 to 972)					
	Zabeau, M. and Stanssens, P.					
	Diagnostic sequencing by a combination of					

mass spectrometry
Patent: WO 0066771-A 3 09-NOV-2000;
Methexis N.V. (BE)
Location/Qualifiers
1..972
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="synthetic-pGEM3-2f(+) derived nucleotide"
BASE COUNT 217 a 272 c 261 g 222 t
ORIGIN

Query Match 98.0%; Score 723; DB 6; Length 972;
Best Local Similarity 100.0%; Pred. No. 1.7e-193;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GGGCGAATTCGAGCTCGGTACCGGGGATCCTCTAGAGTGCAGCTGCAGGATGCAAGCT 75
Db 1 GGGCGAATTCGAGCTCGGTACCGGGGATCCTCTAGAGTGCAGCTGCAGGATGCAAGCT 60

QY 76 TGAGTATTCTATAGTGTACCTTAATAGCTTGGCGTAATCATGTTTCCTG 135
Db 61 TGAGTATTCTATAGTGTACCTTAATAGCTTGGCGTAATCATGTTTCCTG 120

QY 136 TGTGAATTTGTTATCGGCTCACAATTCACACACATAGGCGGGAAGCAATAAGTGA 195
Db 121 TGTGAATTTGTTATCGGCTCACAATTCACACACATAGGCGGGAAGCAATAAGTGA 180

QY 196 AAGCTTGGGCTGCTAATGAGTGAGCTAACTCACATTAATTCGTTGGCTCACTGCCG 255
Db 181 AAGCTTGGGCTGCTAATGAGTGAGCTAACTCACATTAATTCGTTGGCTCACTGCCG 240

QY 256 CTTTCAGTCGGGAACCTGTGTCGACGTGATTAATGATCGCCCAACCGCGGGA 315
Db 241 CTTTCAGTCGGGAACCTGTGTCGACGTGATTAATGATCGCCCAACCGCGGGA 300

QY 316 GAGGCGTTTGGCTATTGGCGCTCTTCGCTTCTCTGCTCACTGCTGCTGCTGCG 375
Db 301 GAGGCGTTTGGCTATTGGCGCTCTTCGCTTCTCTGCTCACTGCTGCTGCTGCG 360

QY 376 TCGTTCGCTGCGGCGAGCGGTATACGTCACTCAAGGCGGTATACGGTTATCCACAG 435
Db 361 TCGTTCGCTGCGGCGAGCGGTATACGTCACTCAAGGCGGTATACGGTTATCCACAG 420

QY 436 AATCAGGGGTAACGAGGGAACATGTGACAAAGCCAGCAAAAGGCGGCAAGC 495
Db 421 AATCAGGGGTAACGAGGGAACATGTGACAAAGCCAGCAAAAGGCGGCAAGC 480

QY 496 GTAAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 555
Db 481 GTAAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

QY 556 AAATCAGCGCTCAAGTCAGAGTGGGGAACCCGACAGACTATAAGATACCGCGCT 615
Db 541 AAATCAGCGCTCAAGTCAGAGTGGGGAACCCGACAGACTATAAGATACCGCGCT 600

QY 616 TTCCCTCGAAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 675
Db 601 TTCCCTCGAAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660

QY 676 TGTCGCGCTTTCCTCCTTCGGAAGCGTGGCGCTTTCATAGCTACGCTGTAGGTATC 735
Db 661 TGTCGCGCTTTCCTCCTTCGGAAGCGTGGCGCTTTCATAGCTACGCTGTAGGTATC 720

QY 736 TCA 738
Db 721 TCA 723

RESULT 4
CVP3EM3Z
LOCUS
DEFINITION Cloning vector pGEM-3Z, circular SYN 18-FEB-2000

ACCESSION X65304
VERSION 3
KEYWORDS beta-lactamase; bla gene; cloning vector; lacZ gene; multiple cloning site; promoter.
SOURCE Cloning vector pGEM-3Z.
ORGANISM Cloning vector pGEM-3Z.
REFERENCE 1 (bases 1 to 2743) Technical Services, Promega Corporation, Direct Submission
AUTHORS Submitted (23-MAR-1992) Technical Services, Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
TITLE revised by [2]
JOURNAL Technical Services.
REMARK 2 (bases 1 to 2743)
AUTHORS Technical Services.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-1993) Technical Services, Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
REMARK 3 (bases 1 to 2743)
AUTHORS revised by [3]
JOURNAL Technical Services.
REMARK 4 (bases 1 to 2743)
AUTHORS Direct Submission
JOURNAL Submitted (26-JAN-2000) Technical Services, Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
REMARK 5 (bases 1 to 2743)
AUTHORS Technical Services.
TITLE Direct Submission
JOURNAL Submitted (18-FEB-2000) Technical Services, Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
COMMENT On Feb 21, 2000 this sequence version replaced g1:6782312. See X65300-X65335 for related vector sequences
This vector can be obtained from Promega Corporation, Madison, WI. Call one of the following numbers for order or technical information:
Order or Technical 800-356-9526
In Wisconsin 800-356-9526
Outside U.S. 608-274-4330.

FEATURES
source
1..2743
/organism="Cloning vector pGEM-3Z"
/db_xref="taxon:90110"
misc_feature
1
/note="T7 transcription initiation site"
misc_feature
5..61
/note="multiple cloning sites"
promoter
67..86
/note="SP6 promoter"
misc_feature
69
/note="SP6 transcription initiation site"
primer_bind
complement(104..120)
/note="pUC/M13 reverse sequencing primer"
misc_feature
complement(106..108)
/gene="lacZ"
/note="lacZ start codon"
gene
complement(106..108)
/gene="lacZ"
misc_feature
128..144
/note="lac operator"
gene
complement(1265..2125)
/gene="bla"
CDS
complement(1265..2125)
/gene="bla"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="CAA46396.2"
/db_xref="GI:6782313"
/translation="MSIOHFRVALIPFFAFCPLVFAHPETLVKVKDAEDQLGARVGY
I ELDLNSGRKILESFRPEERFPMWTFKVLCCGNVLSRIDAGQQLGRRTHYSONDLVE
YSPVTEKHLDITGDRETCIAATMSDNTAANLLLTITGPKELAFALHMGDHYTRL
DRWEPELNEAIIPDERDITPMVAMATTLRLKLTGLLTASRQQLIDHMEADKVAGPL
LRSLPAGFWFIADKSGAGERSGRGIIAALGPDGKPSRIIVYITGSGATMDERNRQIA

misc_feature join(2561..2724,94..323)
 primer_bind /note="lac operon sequence"
 promoter /note="pUC/M13 forward sequencing primer"
 join(2727..2743,1..3)
 /note="T7 promoter"
 BASE COUNT 685 a 685 c 696 g 77 t
 ORIGIN

Query Match 98.0%; Score 7 DB 12; length 2743;
 Best Local Similarity 100.0%; Pred. 1.73-1
 Matches 723; Conservative 0; Mismatch: 0; indels 0; Gaps 0;
 QY 16 GGGCGAATTCGAGCTCGGTACCGGGGATCCCTCTAGAGTCGAGCTCGAGCATGCAAGCT 75
 Db 1 GGGCGAATTCGAGCTCGGTACCGGGGATCCCTCTAGAGTCGAGCTCGAGCATGCAAGCT 75
 QY 76 TGAGTATTCATATGATGTACCTTAATAGCTTGGCGTAATCATGTCATAGCTGTTTCCTG 135
 Db 61 TGAGTATTCATATGATGTACCTTAATAGCTTGGCGTAATCATGTCATAGCTGTTTCCTG 135
 QY 136 TGTGAATTCATATGATGTACCTTAATAGCTTGGCGTAATCATGTCATAGCTGTTTCCTG 120
 Db 121 TGTGAATTCATATGATGTACCTTAATAGCTTGGCGTAATCATGTCATAGCTGTTTCCTG 120
 QY 196 AAGCTTGGGTGCTTAATGAGTACCTTAATGAGTACCTTAATGAGTACCTTAATGAGTACCT 255
 Db 181 AAGCTTGGGTGCTTAATGAGTACCTTAATGAGTACCTTAATGAGTACCTTAATGAGTACCT 255
 QY 256 CTTTCCAGTCGGGAAACCTGCTGCGGAGTGCATTAATGAGTGCATTAATGAGTGCATTAAT 315
 Db 241 CTTTCCAGTCGGGAAACCTGCTGCGGAGTGCATTAATGAGTGCATTAATGAGTGCATTAAT 315
 QY 316 GAGCGCGTTTGGCGTATTTGGCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 375
 Db 301 GAGCGCGTTTGGCGTATTTGGCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 375
 QY 376 TCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAGCGGTAAATGAGTAAATGAGTAAAT 435
 Db 361 TCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAGCGGTAAATGAGTAAATGAGTAAAT 435
 QY 436 AATCAGGGATACGCGAGGAGCAATGATGAGCAAAAGCGGAGCAAAAGCGGAGCAAAAGCG 495
 Db 421 AATCAGGGATACGCGAGGAGCAATGATGAGCAAAAGCGGAGCAAAAGCGGAGCAAAAGCG 495
 QY 496 GTAAAAGCGCGGTGCTGCGGTTTCCATAGGCTCCGCGGCGGCGGCGGCGGCGGCGGCG 555
 Db 481 GTAAAAGCGCGGTGCTGCGGTTTCCATAGGCTCCGCGGCGGCGGCGGCGGCGGCGGCG 555
 QY 556 AAATTCGAGCTCAAGTCAGAGTGGCGAAACCGGAGGAGTAAATGAGTAAATGAGTAAAT 615
 Db 541 AAATTCGAGCTCAAGTCAGAGTGGCGAAACCGGAGGAGTAAATGAGTAAATGAGTAAAT 615
 QY 616 TTCCCGCTTGAAGCTCCCTGCGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 675
 Db 601 TTCCCGCTTGAAGCTCCCTGCGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 675
 QY 676 TGTCGGCTTTCCTCCCTTCGGAAGCGTGGCGCTTTCCTCATAGCTCAGCTGTTAGGTATC 735
 Db 661 TGTCGGCTTTCCTCCCTTCGGAAGCGTGGCGCTTTCCTCATAGCTCAGCTGTTAGGTATC 735
 QY 736 TCA 738
 Db 721 TCA 723

RESULT 5
 CVMEM32FM
 LOCUS
 DEFINITION
 ACCESSION X65307
 3197 bp DNA circular SYN 24-MAY-2000
 Cloning vector pGEM-3zf(-).

X65307.2 GI:8216946
 beta-lactamase; bla gene; cloning vector; lac2 gene; multiple
 cloning site; phage fl region; promoter.
 Cloning vector pGEM-3zf(-).
 Cloning vector pGEM-3zf(-)
 artificial sequence; vectors.
 1 (bases 1 to 3197)
 Solomon, L.R., Massom, L.R. and Jarrett, H.W.
 Enzymatic syntheses of DNA-silicas using DNA polymerase
 Anal Biochem. 203 (1), 58-69 (1992)
 92398067
 2 (bases 1 to 3197)
 Technical Services.
 Direct Submission
 Submitted (23-MAR-1992) Technical Services, Promega Corporation,
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
 revised by [3]
 3 (bases 1 to 3197)
 Technical Services.
 Direct Submission
 Submitted (28-MAY-1993) Technical Services, Promega Corporation,
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
 revised by [4]
 4 (bases 1 to 3197)
 Technical Services.
 Direct Submission
 Submitted (22-MAY-2000) Technical Services, Promega Corporation,
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
 On Jun 3, 2000 this sequence version replaced gi:58172.
 See X65300-X65335 for related vector sequences
 This vector can be obtained from Promega Corporation, Madison, WI.
 Call one of the following numbers for order or technical
 information:
 Order or Technical 800-356-9526
 In Wisconsin 800-356-9526
 Outside U.S. 608-274-4330.
 Location/Qualifiers
 1..3197
 /organism="Cloning vector pGEM-3zf(-)"
 /db_xref="taxon:90113"
 misc_feature
 5..61
 /note="T7 transcription initiation site"
 misc_feature
 67..86
 /note="multiple cloning sites"
 promoter
 69
 /note="SP6 promoter"
 misc_feature
 94..323
 /note="SP6 transcription initiation site"
 primer_bind
 complement(104..120)
 /note="pUC/M13 reverse sequencing primer"
 gene
 complement(106..108)
 /gene="lac2"
 misc_feature
 complement(106..108)
 /gene="lac2"
 gene
 complement(1265..2125)
 /note="lac2 start codon"
 complement(1265..2125)
 /gene="bla"
 CDS
 complement(1265..2125)
 /gene="bla"
 /codon_start=1
 /transl_table=1
 /product="beta-lactamase"
 /protein_id="CAA46399.2"
 /db_xref="GI:8216947"
 /translation="MSIQHFRVALIPFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
 IELDLSKILDSFRPERPMNTFKVLGCVLSRIDAGQEQLGRRIHYQNDELVE
 YSPVTEKHLDGKTVRELCSNAITSDNTANLLLTIGGKELTAFTHNMGDHTVRL
 DRPEPELNEAIPNDERDTTPVAMATTLRKLLTGLLTLASRQQLIDWMEADKVGAPL
 LRSALPAGWFIADKSGAGERSGIIAALGPDGKPSRIVVIYTTGSGATMDERNRQIA
 EIGASLIKHW"
 2562..3017
 misc_feature

misc_feature /note="phage f1 region"
 primer_bind 3018..3178
 promoter /note="lac operon sequence"
 /note="pUC/M13 forward sequencing primer"
 join(3181..3197,1..3)
 /note="T7 promoter"
 781 a 806 c 796 g 814 t

BASE COUNT

ORIGIN

Query Match 98.0%; Score 723; DB 12; Length 3197;
 Best Local Similarity 100.0%; Pred. No. 1.7e-193;
 Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GGCGAATTCAGCTCGGTACCGGGATCCTCTAGAGTCAGCTCGAGGATCAAGCT 75
 DB 1 GGGCGAATTCAGCTCGGTACCGGGATCCTCTAGAGTCAGCTCGAGGATCAAGCT 60
 QY 76 TGAGTATTCATAGTGTACCTAAATAGCTTGGCGTAATCATGTGTATAGCTGTTCTG 135
 DB 61 TGAGTATTCATAGTGTACCTAAATAGCTTGGCGTAATCATGTGTATAGCTGTTCTG 120
 QY 136 TGTGAATTCGTTATCCGCTCACATTCACACACATACGAGCGGAGCATATAAGTGA 195
 DB 121 TGTGAATTCGTTATCCGCTCACATTCACACACATACGAGCGGAGCATATAAGTGA 180
 QY 196 AAGCTGGGTGCTTAAGTAGTAGTCACTCAATTAATGCTGCTGCTGCTGCTGCTG 255
 DB 181 AAGCTGGGTGCTTAAGTAGTAGTCACTCAATTAATGCTGCTGCTGCTGCTGCTG 240
 QY 256 CTTTCCAGTCGGGAACCTGTGCGCAGCTGCAATTAATGAATCGCCGCGGGGA 315
 DB 241 CTTTCCAGTCGGGAACCTGTGCGCAGCTGCAATTAATGAATCGCCGCGGGGA 300
 QY 316 GAGCGGTTTTCGCTATTTGGCGCTTCTCCGCTTCTCCGCTCACTGCTGCTGCTGCTG 375
 DB 301 GAGCGGTTTTCGCTATTTGGCGCTTCTCCGCTTCTCCGCTCACTGCTGCTGCTGCTG 360
 QY 376 TCGTTCGGCTCGCGGAGCGGTATCAGTCACTCAAGGCGGTATACGGTTATCCACAG 435
 DB 361 TCGTTCGGCTCGCGGAGCGGTATCAGTCACTCAAGGCGGTATACGGTTATCCACAG 420
 QY 436 AATCAGGGGTAACCGAGGAAGAACATGTGAGCAAGGCGCAGCAAGGCGGAGAAC 495
 DB 421 AATCAGGGGTAACCGAGGAAGAACATGTGAGCAAGGCGCAGCAAGGCGGAGAAC 480
 QY 496 GTAAAAAGGCGGCTGCTGCGCTTTTTCATAGGCTCCGCGGCTGCTGCTGCTGCTGCTG 555
 DB 481 GTAAAAAGGCGGCTGCTGCGCTTTTTCATAGGCTCCGCGGCTGCTGCTGCTGCTGCTG 540
 QY 556 AAATTCGAGCTCAAGTCAGAGTGGCGAAACCGCAGAGGACTATAAGATACCGGCT 615
 DB 541 AAATTCGAGCTCAAGTCAGAGTGGCGAAACCGCAGAGGACTATAAGATACCGGCT 600
 QY 616 TTCCCGCTGGAAGCTCCGCTGCGCTCTCTGTTTCCGAGCTGCGGCTTACCGGATAC 675
 DB 601 TTCCCGCTGGAAGCTCCGCTGCGCTCTCTGTTTCCGAGCTGCGGCTTACCGGATAC 660
 QY 676 TGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTTCATAGCTCACTGCTAGGTATC 735
 DB 661 TGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTTCATAGCTCACTGCTAGGTATC 720
 QY 736 TCA 738
 DB 721 TCA 723

RESULT 6
 CVP32P CVP32P 3197 bp DNA circular SYN 16-JUL-1999
 LOCUS
 DEFINITION Cloning vector pGEM-3zf(+).
 ACCESSION X65306

X65306.2 GI:5531232
 Cloning vector; multiple cloning site; phage f1 region; promoter.
 Cloning vector pGEM-3zf(+).
 Cloning vector pGEM-3zf(+).
 artificial sequence; vectors.
 1 (bases 1 to 3197)
 Solomon, L.R., Massom, L.R. and Jarrett, H.W.
 Enzymatic syntheses of DNA-silicas using DNA polymerase
 Anal. Biochem. 203 (1), 58-69 (1992)
 92398067
 2 (bases 1 to 3197)
 Technical Services.
 Direct Submission
 Submitted (23-MAR-1992) Technical Services, Promega Corporation,
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
 revised by [3]
 3 (bases 1 to 3197)
 Technical Services.
 Direct Submission
 Submitted (28-MAY-1993) Technical Services, Promega Corporation,
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
 revised by [4]
 4 (bases 1 to 3197)
 Kenefick, K.
 Direct Submission
 Submitted (14-JUL-1999) Kenefick K., Technical Services, Promega
 Corporation, 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
 On Jul 19 1999 this sequence version replaced gi:58203.
 See X65306-X65335 for related vector sequences
 This vector can be obtained from Promega Corporation, Madison, WI.
 Call one of the following numbers for order or technical
 information:
 Order or Technical 800-356-9526
 In Wisconsin 800-356-9526
 Outside U.S. 608-274-4330.
 Location/Qualifiers
 1..3197
 /organism="Cloning vector pGEM-3zf(+)"
 /db_xref="taxon:90112"
 misc_feature 1 /note="T7 transcription initiation site"
 misc_feature 5..61 /note="multiple cloning sites"
 promoter 67..86 /note="SP6 promoter"
 misc_feature 69 /note="SP6 transcription initiation site"
 misc_feature 94..323 /note="lac operon sequence"
 complement(106..108)
 /gene="lacZ"
 /note="lacZ start codon"
 complement(106..108)
 /gene="lacZ"
 complement(1265..2125)
 /gene="bla"
 complement(1265..2125)
 /gene="bla"
 /codon_start=1
 /transl_table=11
 /product="beta-lactamase"
 /protein_id="CAA46398.2"
 /db_xref="GI:5531233"
 /translation="MSIQHFRVALIPFAAFCLPVFAHPETLVKVKDAEDOLGARVGY
 IELDLSNGKILSEFRPEERPMSTFKVLICGLAVLRIDAGQQLGRRIRHYSDNLVE
 YSPVTEKHLTDGMTVRELCSAALTMSDNTAANLLLTIGGPKELTAFIHNMGKDVTRL
 DRWPELNEAIFNDRDTPVAMATTLKLLTGELLTLASRQLIDWMEADKVAQPL
 LRSALPAGFIADKSGAGERSGIITAAIGDPKSRIVVITTTGSGATWDERNRQIA
 EICASLIKHW"
 2562..3017
 /note="phage f1 region"
 3018..3178
 /note="lac operon sequence"
 misc_feature 2562..3017
 misc_feature 3018..3178

```

promoter      join(3181...3197,3)
BASE COUNT    822 a      785 c      817 g      773 t
ORIGIN

Query Match      98.0%; Score 723; DB 12; Length 3197;
Best Local Similarity 100.0%; Pred. No. 1.7e-193;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GGCGCAATTCGAGCTCGGTACCGGGGATCCTCTAGAGTCGAGCTGCAGCGCATGCAAGCT 75
DB 1 GGCGCAATTCGAGCTCGGTACCGGGGATCCTCTAGAGTCGAGCTGCAGCGCATGCAAGCT 60
QY 76 TGAGTATTCATAGTCTCACTAAATAGCTTGGCGTAATCATGTCATAGCTGTTCCTG 135
DB 61 TGAGTATTCATAGTCTCACTAAATAGCTTGGCGTAATCATGTCATAGCTGTTCCTG 120
QY 136 TGTGAATTTGTTATCCGCTCACAAATCCACACACATACAGCGCGGAAGCATAAAGTGA 195
DB 121 TGTGAATTTGTTATCCGCTCACAAATCCACACACATACAGCGCGGAAGCATAAAGTGA 180
QY 196 AAGCCTGGGGTGCTTAATGAGTGAGCTAACTACATTAATTTGCGTTCGCTCACTGCCG 255
DB 181 AAGCCTGGGGTGCTTAATGAGTGAGCTAACTACATTAATTTGCGTTCGCTCACTGCCG 240
QY 256 CTTTCAGTCGGGAACCTGTGTCGCGCTGCTGCAATTAATGAATCGGCCAACGCGCGGA 315
DB 241 CTTTCAGTCGGGAACCTGTGTCGCGCTGCTGCAATTAATGAATCGGCCAACGCGCGGA 300
QY 316 GAGGCGTTTGGTATTGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTCGCTCGG 375
DB 301 GAGGCGTTTGGTATTGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTCGCTCGG 360
QY 376 TCGTTTCGCTGCGGAGCGGTATCAGCTCACTCAAGGCGGTAAATACGGTTATCCACAG 435
DB 361 TCGTTTCGCTGCGGAGCGGTATCAGCTCACTCAAGGCGGTAAATACGGTTATCCACAG 420
QY 436 AATCAGGGGATACCGCAGGAAGACATGTGAGCAAAAGCGCAGCAAAAGCGCAGAAC 495
DB 421 AATCAGGGGATACCGCAGGAAGACATGTGAGCAAAAGCGCAGCAAAAGCGCAGAAC 480
QY 496 GTAAAGGCGCGTGTGCTGCGCTTTTCCATAGGCTCGCGCCCTCGACGAGCATCA 555
DB 481 GTAAAGGCGCGTGTGCTGCGCTTTTCCATAGGCTCGCGCCCTCGACGAGCATCA 540
QY 556 AAATTCGAGCTCAAGTCAGAGGTGGCGAAACCGCAGGACTATAAAGATACCGAGCGT 615
DB 541 AAATTCGAGCTCAAGTCAGAGGTGGCGAAACCGCAGGACTATAAAGATACCGAGCGT 600
QY 616 TTCCCGCTGGAGCTCCCTCGTCCGCTCTCTCTGTTCCGACCTCGCGCTTACCGGATAC 675
DB 601 TTCCCGCTGGAGCTCCCTCGTCCGCTCTCTCTGTTCCGACCTCGCGCTTACCGGATAC 660
QY 676 TGTCCGCTTTCTCCCTCGGAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGTATC 735
DB 661 TGTCCGCTTTCTCCCTCGGAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGTATC 720
QY 736 TCA 738
DB 721 TCA 723

RESULT 7
SYNBLUEV      SYNBLUEV      2746 bp      DNA      circular SYN 16-MAR-2000
LOCUS         BlueScribe cloning vector.
DEFINITION    M7811 M7798
ACCESSION     M7811.1 GI:208035
VERSION       M7811.1
KEYWORDS      Cloning vector BlueScribe.
SOURCE        Cloning vector BlueScribe
              artificial sequence; vectors.

```

```

1 (sites)
Short J.M., Fernandez J.M., Sorge J.A. and Huse W.D.
Lambda ZAP-A bacteriophage lambda expression vector with in vivo
excision properties
Nucleic Acids Res. 16, 7583-7600 (1988)
88319944
2 (bases 1 to 2746)
Gilbert W.
Obtained from VecBase 3.0
Unpublished
These data and their annotation were supplied to GenBank by Will
Gilbert under the auspices of the GenBank Curator Program.
BlueScribe - Cloning vector
ENTRY BLUE
- Cloning vector
#TYPE DNA CIRCULAR TITLE BlueScribe
DATE
#sequence 03-MAR-1986
#sequence 16-DEC-1986
#sequence 02-FEB-1987
#sequence 04-MAR-1987
ACCESSION VB0042
SOURCE artificial
REFERENCE
#number 1
#authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorge
J.
#Journal Gene (1987) in press
#citation sequence data from Stratagene
#comment sequence correction according to Stratagene COMMENT
Obtained 28-JAN-1987 from Stratagene on floppy disc.
Revised 2-FEB-1986 by F. Pfeiffer:
1122/3 'AT' to 'TA' to match revised sequence of pBR322
Revised 4-MAR-1987 to match sequence of pUC19 on request
of Stratagene
COMMENT
The stand shown corresponds to pUC19c.
As in the published sequence of pUC19c, The M13mp19 lacZ region
is on the complementary strand.
KEYWORDS
CROSSREFERENCE
#parent
VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3
#offspring
VecBase(3):BlueM13p, VecBase(3):BlueM13m
PARENT
Features of BlueScribe (2746 bp)
residue source
1- 400 684- 281 (c) pUC19
402- 421 1- 20 T7 promoter
423- 479 1- 57 M13mp19/pUC19-polylinker
485- 506 20- 1 (c) T3 promoter
508- 744 237- 1 (c) pUC19
745-2746 2686- 685 (c) pUC19
Conflict (cfl) and Mutations (mut): none
PARENT
Features of BlueScribe (2746 bp)
residue source
1- 400 1- 400 pUC19c
402- 421 1- 20 T7 promoter
423- 479 57- 1 (c) M13mp19/pUC19-polylinker
485- 506 20- 1 (c) T3 promoter
513-2746 453-2686 pUC19c
Conflict (cfl) and Mutations (mut): none
FEATURE
1689-2477 789-1 (c) Ap-R; b-lactamase
POLYLINKER HindIII-SphI-PstI-SalI-XbaI-BamHI-SmaI-KpnI-SacI-EcoRI
SELECTION
#resistance Ap
#indicator beta-galactosidase
SUMMARY Blue #length 2746 #checksum 2145.
1..2746
Location/Qualifiers
/organism="Cloning vector BlueScribe"
/db_xref="taxon:118306"
FEATURES
source

```

```

/focus
/note="BlueScribe cloning vector"
1. .400
/organism="Cloning vector pUC19"
/db_xref="taxon:31851"
402. .421
/organism="Enterobacteria phage T7"
/note="Bacteriophage T7 promoter"
/db_xref="taxon:10760"
423. .479
/organism="Cloning vector pUC19"
/db_xref="taxon:31851"
485. .506
/organism="Bacteriophage T3"
/note="Bacteriophage T3 promoter"
/db_xref="taxon:10759"
508. .744
/organism="Cloning vector pUC19"
/db_xref="taxon:31851"
745. .2746
/organism="Cloning vector pUC19"
/db_xref="taxon:31851"
BASE COUNT 680 a 686 c 700 g 680 t
ORIGIN

```

Query Match 95.68; Score 705.8; DB 12; Length 2746;
 Best Local Similarity 98.08; Pred. No. 1.2e-188;
 Matches 726; Conservative 0; Mismatches 12; Indels 3; Gaps 1;

```

QY 1 ATAGGACTCACTATAGGCGGAATTCGAGCTCGGTACCGGGGATCCTCTAGAGTCGACCT 60
DB 404 ATAGGACTCACTATAGGCGGAATTCGAGCTCGGTACCGGGGATCCTCTAGAGTCGACCT 463
QY 61 GCAGGCGTCAAGCTTGAGTATTCATATAGTGTCACTAAAT---AGCTTGGCGTAAATCAT 117
DB 464 GCAGGCGTCAAGCTTTGTTCCCTTTAGTGGGGTTAATTCGAGCTTGGCGTAAATCAT 523
QY 118 GGTATAGCTGTTTCCTGTGTGAATTTGTTATCCGTCACAAATCCACACATACGAG 177
DB 524 GGTATAGCTGTTTCCTGTGTGAATTTGTTATCCGTCACAAATCCACACATACGAG 583
QY 178 CCGAAGCATAAAGTGAAGCCCTGGGTGCTTAATGATGAGTAACTCACTAAATG 237
DB 584 CCGAAGCATAAAGTGAAGCCCTGGGTGCTTAATGATGAGTAACTCACTAAATG 643
QY 238 CGTTGCGCTCACTGCCCGCTTTCACAGTCGGGAAACCTGTGTCGAGCTGCATTAATGAA 297
DB 644 CGTTGCGCTCACTGCCCGCTTTCACAGTCGGGAAACCTGTGTCGAGCTGCATTAATGAA 703
QY 298 TCGGCCAAGCGCGGGGAGAGCGGTTTTCGTTATTTGGGCGCTTTCGCTTCCCTCGCTCA 357
DB 704 TCGGCCAAGCGCGGGGAGAGCGGTTTTCGTTATTTGGGCGCTTTCGCTTCCCTCGCTCA 763
QY 358 CTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCA 417
DB 764 CTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCA 823
QY 418 TAATACGCTTATCCACAGATACAGGGATACCGAGGAGAAACATGTGAGCAAAAGGCC 477
DB 824 TAATACGCTTATCCACAGATACAGGGATACCGAGGAGAAACATGTGAGCAAAAGGCC 883
QY 478 AGCAAAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 537
DB 884 AGCAAAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 943
QY 538 CCCCTGAGGAGCATCACAAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGGAC 597
DB 944 CCCCTGAGGAGCATCACAAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGGAC 1003
QY 598 TATAAGATACAGCGGTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCTGTTTCCGACCC 657
DB 1004 TATAAGATACAGCGGTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCTGTTTCCGACCC 1063

```

```

QY 658 TGGCGCTTACCGGATACCTGTCGCGCTTTCCTCCCTTCGGGAAGCGTGGCGCTTCTCATA 717
DB 1064 TGGCGCTTACCGGATACCTGTCGCGCTTTCCTCCCTTCGGGAAGCGTGGCGCTTCTCATA 1123
QY 718 GCTCAGCGCTAGGTATCTCA 738
DB 1124 GCTCAGCGCTAGGTATCTCA 1144

RESULT 8
SYNBLM13MV LOCUS 3204 bp DNA circular SYN 26-JUL-1993
DEFINITION BlueScribe M13 Minus cloning vector.
ACCESSION L08782
VERSION L08782.1 GI:310730
KEYWORDS Synthetic construct DNA.
ORGANISM synthetic construct
SOURCE artificial sequence.
REFERENCE 1 (bases 1 to 3204)
AUTHORS Gilbert,W.
TITLE Obtained from VecBase 3.0
JOURNAL Unpublished (1991)
COMMENT These data and their annotation were supplied to GenBank by Will
        Gilbert under the auspices of the GenBank Curator Program.
        BlueScribe M13 Minus - Cloning vector
ENTRY BLUEM13M #TYPE DNA CIRCULAR TITLE Bluescribe
M13 Minus - Cloning vector
DATE 27-MAR-1986
#sequence 16-DEC-1986
#sequence 04-MAR-1987
ACCESSION VB0043
SOURCE artificial
REFERENCE
#number 1
#authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorge J.
        #journal Gene (1987) in press
        #citation Sequence data from StrataGene
        #comment sequence correction according to StrataGene COMMENT
        Obtained 28-JAN-1987 from StrataGene on floppy disc.
        Revised 2-FEB-1986 by F. Pfeiffer:
        1580/1 'AT' to 'TA' to match revised sequence of pBR322
        Revised 4-MAR-1987 to match sequence of pUC19 on request
        of StrataGene
COMMENT
        The stand shown corresponds to pUC19c.
        As in the published sequence of pUC19c, The M13mpl9 lacZ region
        is on the complementary strand.
COMMENT
        Despite its name Blue M13 minus, this vector contains the
        phase fl origin so that the minus strand can be obtained upon
        phase fl superinfection.
KEYWORDS
CROSSREFERENCE
#parent
        VecBase(3):Blue
        VecBase(3):BlueM13p
#parent
        VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,
        GenBank(50):PFI
#brother
        VecBase(3):BlueM13p
#offspring
        VecBase(3):BlueKsm, VecBase(3):BlueSkm
PARENT
        Features of BlueM13m (3204 bp)
        residue source
        1- 238 684- 447 (c) pUC19
        239- 694 5488-5943 phase fl
        695- 858 448- 285 (c) pUC19
        860- 879 1- 20 T7 promoter
        881- 937 57- 1 (c) M13mpl9/pUC19-polylinker

```

D 1402 CCCTGACGAGCATCACAAAAATCGACGGCTCAAGTCAGAGTGGCGAAACCCGACAGGAC 1461


```

Query Match          93.7%; Score 691.8; DB 6; Length 3216;
Best Local Similarity 96.2%; Pred. No. 1.le-184;
Matches 726; Conservative 0; Mismatches 12; Indels 17; Gaps 1;

1 ATACGAGCTCACTATAGGGCGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTGCAGACT 60
|||||
686 ATACGAGCTCACTATAGGGCGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCSAAG 745
|||||

61 GCAGGC-----ATGCAAGCTTGAGTATTCCTATATAGTGTCACTTAAGTAG 103
|
746 CTTCTCGCCCTATAGTGAGTCGTATTACAGCTTGAGTATTCTATAGTGTCACTAAATAG 805
|||||

104 CTTGGCGTAATCATGTTTCATAGCTGTTTCTGCTGAAATGTTATTCGCGCTCACAAATGCC 163
|||||

806 CTTGGCGTAATCATGTTTCATAGCTGTTTCTGCTGTAATTTGTTATCCGCTCACAAATGCC 865
|||||

164 ACACAACATACGAGCGCGGAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTA 223
|||||

866 ACACAACATACGAGCGCGGAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTA 925
|||||

224 ACTCACAATTAATTCGGTTTGGCGTCACTGCCCGGCTTTCCAGTCGGGGAACCTGTCGTGCCA 283
|||||

926 ACTCACAATTAATTCGGTTTGGCGTCACTGCCCGGCTTTCCAGTCGGGGAACCTGTCGTGCCA 985
|||||

284 GTCGCAATTAATGAATCGGCCCAACCGCGGGGAGAGCGGTTTCGCTATTGGGCGCTCTTTC 343
|||||

986 GCTGCAATTAATGAATCGGCCCAACCGCGGGGAGAGCGGTTTCGCTATTGGGCGCTCTTTC 1045
|||||

344 CGCTTTCCTCGGCTCACTGACTCGCTCGGCTCGGTCGCTCGGCTCGCGGAGCGGCTATCAGC 403
|||||

```


Query Match 93.4%; Score 689.4; DB 6; Length 8126;
 Best Local Similarity 99.9%; Pred. No. 5.4e-184;
 Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 48 CTAGAGTGCACCTGCAGGATCAAGCTTCAGTATTTCTATAGTGTCTACCTAAATAGCTTG 107
 Db 4956 CTAGAGTGCACCTGCAGGATCAAGCTTCAGTATTTCTATAGTGTCTACCTAAATAGCTTG 5015

QY 108 GCCTAATCATGGTCATAGCTGTTTCTCTGTGTGAATTTGTTATCCGCTCACAATTCACAC 167
 Db 5016 GCCTAATCATGGTCATAGCTGTTTCTCTGTGTGAATTTGTTATCCGCTCACAATTCACAC 5075

QY 168 ACATAGCAGCCGGAAGCATAAAGTGTAAAGCTTGGGTCCTAATGAGTACGTAACCTC 227
 Db 5076 ACATAGCAGCCGGAAGCATAAAGTGTAAAGCTTGGGTCCTAATGAGTACGTAACCTC 5135

QY 228 ACATTAATTCGGTTCGCTCACTGCCGCTTTTCAGTCGGGAACCTGTGCTGCCAGCTG 287
 Db 5136 ACATTAATTCGGTTCGCTCACTGCCGCTTTTCAGTCGGGAACCTGTGCTGCCAGCTG 5195

QY 288 CATTAATGAATCGCCCAACCGCGGGAGAGCGGTTTGGTATTTGGCGCTCTTCGCT 347
 Db 5196 CATTAATGAATCGCCCAACCGCGGGAGAGCGGTTTGGTATTTGGCGCTCTTCGCT 5255

QY 348 TCCTCGCTCACTGACTCGCTGCGTCTGCTTTCGCTGCGGCGAGCGGTATCAGCTCAC 407
 Db 5256 TCCTCGCTCACTGACTCGCTGCGTCTGCTTTCGCTGCGGCGAGCGGTATCAGCTCAC 5315

QY 408 TCAAGGCGGTATACCGTTTATCCACAGATTCAGGGGATACCGAGGAAGAAACATGTGA 467
 Db 5316 TCAAGGCGGTATACCGTTTATCCACAGATTCAGGGGATACCGAGGAAGAAACATGTGA 5375

QY 468 GCAAAAGCCAGCAAAAGCCAGCAACCGTAAAGGCGCGTTCGCTGGCGTTTTCCT 527
 Db 5376 GCAAAAGCCAGCAAAAGCCAGCAACCGTAAAGGCGCGTTCGCTGGCGTTTTCCT 5435

QY 528 AGGCTCGCCCTGAGCAGCATCACAAAATCGAGCTCAAGTCAGAGTGGCGAAAC 587
 Db 5436 AGGCTCGCCCTGAGCAGCATCACAAAATCGAGCTCAAGTCAGAGTGGCGAAAC 5495

QY 588 CCGAGAGCTATTAAGATACAGCGGTTTCCCTCGGAAGCTCCCTCGTGGCTCTCTCT 647
 Db 5496 CCGAGAGCTATTAAGATACAGCGGTTTCCCTCGGAAGCTCCCTCGTGGCTCTCTCT 5555

QY 648 GTTCCGACCTGCGCTTACCGGATACCTGTCGCTTCTCCCTTCGGGAGCGTGGCG 707
 Db 5556 GTTCCGACCTGCGCTTACCGGATACCTGTCGCTTCTCCCTTCGGGAGCGTGGCG 5615

QY 708 CTTTCTCATAGCTCAGCGTGTAGGTATCTCA 738
 Db 5616 CTTTCTCATAGCTCAGCGTGTAGGTATCTCA 5646

RESULT 13
 CVP4M42
 LOCUS CVP4M42 2746 bp DNA circular SYN 26-JAN-2000
 DEFINITION Cloning vector pGEM-4Z.
 ACCESSION X65305
 VERSION X65305.2 GI:6782314
 KEYWORDS beta-lactamase; bla gene; cloning vector; lacZ gene; multiple cloning site; promoter.
 SOURCE Cloning vector pGEM-4Z.
 ORGANISM artificial sequence; vectors.
 REFERENCE 1 (bases 1 to 2746)
 AUTHORS Technical, Services.
 TITLE Direct Submission
 JOURNAL Submitted (23-MAR-1992) Technical Services, Promega Corporation,
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
 REMARK revised by [2]
 REFERENCE 2 (bases 1 to 2746)

Technical, Services.
 Direct Submission
 Submitted (28-MAY-1993) Technical Services, Promega Corporation,
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
 REMARK revised by [3] (2746)
 REFERENCE 3 (bases 1 to 2746)
 AUTHORS Technical, Services.
 TITLE Direct Submission
 JOURNAL Submitted (26-JAN-2000) Technical Services, Promega Corporation,
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
 COMMENT On Jan 27, 2000 this sequence version replaced gi:58207.
 See X65300-X65335 for related vector sequences
 This vector can be obtained from Promega Corporation, Madison, WI.
 Call one of the following numbers for order or technical information:
 Order or Technical 800-356-9526
 In Wisconsin 800-356-9526
 Outside U.S. 608-274-4330.

FEATURES
 source
 1. .2746
 /location/Qualifiers
 /organism="Cloning vector pGEM-4Z"
 /db_xref="taxon:90111"

misc_feature
 1
 /note="SP6 transcription initiation site"

misc_feature
 7. .63
 /note="multiple cloning sites"

promoter
 68. .87
 /note="T7 promoter"

misc_feature
 70
 /note="T7 transcription initiation site"

misc_feature
 96. .325
 /note="lac operon sequence"

primer_bind
 complement(106. .122)
 /note="pUC/M13 reverse sequencing primer"

gene
 complement(108. .110)
 /gene="lacZ"

misc_feature
 complement(108. .110)
 /gene="lacZ"

misc_feature
 130. .146
 /note="lac start codon"

gene
 complement(1267. .2127)
 /note="lac operator"

CDS
 complement(1267. .2127)
 /gene="bla"

codon_start=1
 /transl_table=11
 /product="beta-lactamase"

protein_id="CAA46397.2"
 /db_xref="GI:6782315"

translation="MSIQHFRVALIPFFAFCFLPVFAHPETLVKVKDAEDQLGARVGY
 IELDLSKGKILESFRPEERPMSTFKVLLCGAVLSRIDAGQEQGLRRIRHYSQNDLVE
 YSPVTEKHLDGMITVRELCSAAITMSDNTAANLLTTIGPKELTAFHNMGDHVTPL
 DRWEPELNEAIPNDRDITMPVAMATTLKRLTGLTLLASROQLIDWMEADKRVAGPL
 LRSALPAGWFIADKSGAGERSGRIIAALGPDGKPSRIIVITVTGSOATMDERNROIA
 EIGASLIKHW"

misc_feature
 2563. .2726
 /note="lac operon sequence"

primer_bind
 2686. .2702
 /note="pUC/M13 forward sequencing primer"

promoter
 join(2730. .2746,1..3)
 /note="SP6 promoter"

BASE COUNT 684 a 684 c 699 g 679 t
 ORIGIN

Query Match 92.8%; Score 684.8; DB 12; Length 2746;
 Best Local Similarity 96.9%; Pred. No. 1.1e-182;
 Matches 698; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 19 CGAATTCGAGCTCGGTACCGGGGATCCTCTAGAGTGCAGCTGCAGGCATGCAAGCTTGA 78
 Db 6 CGAATTCGAGCTCGGTACCGGGGATCCTCTAGAGTGCAGCTGCAGGCATGCAAGCTTGT 65


```

QY 497 TAAAAAGCGCGTGTCTGGCGTTTTCATAGGCTCCGCCCTCCGACGAGCATCAAA 556
Db 1352 TAAAAAGCGCGTGTCTGGCGTTTTCATAGGCTCCGCCCTCCGACGAGCATCAAA 1411
QY 557 AATCGACGCTCAAGTCAGAGTGGCGAAACCGACGAGGACTATAAGATACCAGCGTT 616
Db 1412 AATCGACGCTCAAGTCAGAGTGGCGAAACCGACGAGGACTATAAGATACCAGCGTT 1471
QY 617 TCCCTCGGAAGCTCCCTGGCGCTCTCTGTTCGACGACCTCCGCTTACCAGTACCT 676
Db 1472 TCCCTCGGAAGCTCCCTGGCGCTCTCTGTTCGACGACCTCCGCTTACCAGTACCT 1531
QY 677 GTCCGCTTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCT 736
Db 1532 GTCCGCTTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCT 1591
QY 737 CA 738
Db 1592 CA 1593

RESULT 15
AF134573
LOCUS AF134573 3946 bp DNA linear SYN 03-MAY-1999
DEFINITION Cloning vector p34S-Sm3, complete sequence.
ACCESSION AF134573
VERSION AF134573.1 GI:4731624
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Dennis.J.J. and Zylstra.G.J.
TITLE Construction of a modified Sm/Sp resistance cassette
JOURNAL Unpublished
AUTHORS Dennis.J.J. and Zylstra.G.J.
TITLE Direct Submission
JOURNAL Submitted (12-MAR-1999) Biotechnology Center for Agriculture and
the Environment, Cook College, Rutgers University, 59 Dudley Rd.,
New Brunswick, NJ 08901-8520, USA
FEATURES
source
1..3946
/organism="Cloning vector p34S-Sm3"
/db_xref="taxon:92792"
/note="antibiotic resistance cassette plasmid"
misc_feature
1..51
/note="multiple cloning site"
gene
418..1209
/gene="AAD"
CDS
418..1209
/gene="AAD"
/note="aminoglycoside adenyltransferase 3' end"
/codon_start=1
/transl_table=11
/product="streptomycin/spectinomycin adenyltransferase"
/protein_id="AAD28532.1"
/db_xref="GI:4731625"
/transl_table=11
/translation="MREAVIARVSTQLSEVGVIERHLEPTLLAVHLYGSVDGGLKP
HSDIDLVTVRDETRRALINDLETASPGSEILRAVEVTVVHDDIIPWRYP
AKREIFQWQRNDILAGIFEPATIDILAILTKAREHSVALVGPAAELFPVPEQ
DLFEALNETLWNSPPDWAGDERNVLTLSRIWYSAVTGKIAPKDAADWAWERLPA
QYQVILEARQAYLQEEEDRLASRADOLEEFVHYVKEITKVVGK"
misc_feature
1210..1260
/note="multiple cloning site"
gene
complement(2464..3324)
/gene="bla"
CDS
complement(2464..3324)
/gene="bla"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AAD28533.1"

```

```

/db_xref="GI:4731626"
/translation="MSIQHFRVALIPFAARCLPVFAHPETLVKVKDAEDQLGARVGY
IELDNKGKILDFSRPREFPMSTFRVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
YDPTTEKHLTDGMTVRELCSAAITMSDNTAANLLLTIGGKELTAFTHHNGDHLVRL
DRWEPELNEAIPNDERDTTPMPAMPTTRIKLTLLTGLLTSLASRQQLIDWMEADKVGPL
LRSLAPAGWEIADKSGAGERSGRIIAALGPDGKPSRIIVTVITTTGSGQATMDERNQIA
EIGASLIKHN"
BASE COUNT 971 a 983 c 1047 g 945 t
ORIGIN
Query Match 92.0%; Score 679; DB 12; Length 3946;
Best Local Similarity 95.9%; Pred. No. 4.7e-181;
Matches 697; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 12 TATAGGCGGAATTCAGCTCGGTACCCGGGGATCCTCTAGAGTCACCTGCGAGGATGCA 71
Db 1196 TAGTCGGCAATAAGAGCTCGGTACCCGGGGATCCTCTAGAGTCACCTGCGAGGATGCA 1255
QY 72 AGCTTGAGTATCTATAGTGTACCTAAATAGCTTGGCGTAAATCATGCTAGCTGTTT 131
Db 1256 AGCTTGCTCCCTATAGTGTAGTGTATAGAGCTTGGCGTAAATCATGCTAGCTGTTT 1315
QY 132 CCTGTGTAAATTTGATCCCTCACAAATTCACACAAACATACGAGCGGAAAGCATAAAG 191
Db 1316 CCTGTGTAAATTTGATCCCTCACAAATTCACACAAACATACGAGCGGAAAGCATAAAG 1375
QY 192 TGTAAAGCCTGGGTGCTTAATAGTGTAGCTAATCAATTAATTCGCTTGGCTCAGCTG 251
Db 1376 TGTAAAGCCTGGGTGCTTAATAGTGTAGCTAATCAATTAATTCGCTTGGCTCAGCTG 1435
QY 252 CCCGCTTTCCAGTCGGGAAACCTGTCTGCCAGCTGCATTAAATGAATGGCCAAACGCGG 311
Db 1436 CCCGCTTTCCAGTCGGGAAACCTGTCTGCCAGCTGCATTAAATGAATGGCCAAACGCGG 1495
QY 312 GGGAGAGCGCGTTTTCGCTATTGGCGCTCTTTCGCGCTTCTCGCTCACTGCTCGCTCGCG 371
Db 1496 GGGAGAGCGCGTTTTCGCTATTGGCGCTCTTTCGCGCTTCTCGCTCACTGCTCGCTCGCG 1555
QY 372 TCGTGTCTGGCTGGCGGAGCGGTATCAGCTCACTCAAAAGCGGTATACGGTTATGCC 431
Db 1556 TCGTGTCTGGCTGGCGGAGCGGTATCAGCTCACTCAAAAGCGGTATACGGTTATGCC 1615
QY 432 ACAGAAATCAGGGGATACGAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGG 491
Db 1616 ACAGAAATCAGGGGATACGAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGG 1675
QY 492 AACCGTAAAAAGCGCGCTTTCGCTGGCGTTTTCATAGGCTCCGCCCTCCCTGACGAGCAT 551
Db 1676 AACCGTAAAAAGCGCGCTTTCGCTGGCGTTTTCATAGGCTCCGCCCTCCCTGACGAGCAT 1735
QY 552 CACAAAATCGACGCTCAAGTCAGAGTGGCGAAACCGACGAGGACTATAAGATACCAAG 611
Db 1736 CACAAAATCGACGCTCAAGTCAGAGTGGCGAAACCGACGAGGACTATAAGATACCAAG 1795
QY 612 GCGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTGTTCGACGACCTCCGCTTACCCTG 671
Db 1796 GCGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTGTTCGACGACCTCCGCTTACCCTG 1855
QY 672 TACCTGTCCGCTTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGG 731
Db 1856 TACCTGTCCGCTTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGG 1915
QY 732 TATCTCA 738
Db 1916 TATCTCA 1922

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2002, 09:37:28 ; Search time 365.16 Seconds

(without alignments)
3469.939 Million cell updates/sec

Title: US-10-014-743-3
Perfect score: 738
Sequence: 1 ATACGACTACTATAGGCG.....CTCAGCGTGTAGTATCTCA 738

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	738	100.0	738	20	AAZ09716
2	723	98.0	972	21	AAAC83524
3	723	98.0	3253	19	AAV43439
4	691.8	93.7	3216	21	AAZ49993
5	691.8	93.7	3216	22	AAZ06382
6	689.4	93.4	8126	21	AAZ29606
7	672.4	91.1	4283	18	AAZ86449
8	672.4	91.1	4283	19	AAZ86610
9	672.4	91.1	4283	21	AAA63237

10	669	90.7	4118	18	AAZ69188
11	669	90.7	4118	18	AAZ69189
12	667.8	90.5	4539	16	AAQ87347
13	667.8	90.5	4539	19	AAV37252
c 14	667	90.4	2571	21	AAZ62824
c 15	667	90.4	2571	24	ABK17089
c 16	667	90.4	3018	20	AAZ29905
17	667	90.4	3968	22	AAZ09981
18	667	90.4	4229	22	AAZ25896
19	667	90.4	4626	22	AAZ09988
20	667	90.4	5919	22	AAZ09980
21	667	90.4	7175	21	AAA26728
22	667	90.4	7177	21	AAA52989
23	666.8	90.4	10930	16	AAQ81226
24	666.8	90.4	10950	16	AAQ81225
25	666	90.2	4454	22	AAZ06386
26	666	90.2	4620	22	AAZ06383
27	666	90.2	4643	22	AAZ06385
28	666	90.2	4701	22	AAZ06387
29	666	90.2	4756	22	AAZ06384
30	665.8	90.2	6688	22	AAZ02333
c 31	665.4	90.2	5115	20	AAZ08719
c 32	665.4	90.2	8320	22	AAZ02322
33	665.4	90.2	8803	22	AAZ02321
34	665.4	90.2	13928	22	AAH77500
c 35	660.4	89.5	10306	15	AAQ73734
c 36	660.4	89.5	10970	15	AAQ73735
c 37	659.4	89.3	9837	18	AAZ96851
c 38	650.8	88.2	4161	24	ABA04699
39	650.4	88.1	742	22	AAH82483
c 40	644.6	87.3	3854	15	AAQ67221
c 41	642.4	87.0	4344	21	AAA53848
c 42	642.4	87.0	5753	21	AAZ53835
c 43	642.4	87.0	5753	22	AAZ18173
c 44	642.4	87.0	8720	22	AAZ18172
c 45	642.4	87.0	10698	22	AAZ18176

ALIGNMENTS

RESULT 1	
AAZ09716	
ID	AAZ09716 standard; DNA; 738 BP.
XX	
AC	AAZ09716;
XX	
DT	15-NOV-1999 (first entry)
XX	
DE	Plasmid pGEM DNA fragment.
XX	
DE	Energy transfer dye; fluorescence; donor; acceptor; excitation energy;
KW	light absorbtion; nucleic acid sequencing; detection; fluorophore; ss.
XX	
OS	Synthetic.
XX	
PN	US5945526-A.
XX	
PD	31-AUG-1999.
XX	
PF	23-MAR-1998; 98US-0046203.
XX	
PR	23-MAR-1998; 98US-0046203.
PR	03-MAY-1996; 96US-0642330.
PR	27-JUN-1996; 96US-0672196.
XX	
PA	(PEKE) PERKIN-ELMER CORP.
XX	
PI	Lee LG, Rosenblum B, Spurgeon SL;
XX	
DR	WPI; 1999-550120/46.
XX	
PT	New energy transfer dye, used in nucleic acid sequencing

Construct pGEM-htr
Construct pGEM-htr
Plasmid pINVI. Sy
pINVI plasmid used
Glyceroldehydro-3-p
Eucaalyptus grandis
Plasmid pGEM (RTM)
pISP70-1MCS constr
Genetic informatio
pHSP-oHoxDS/BH pla
pHSP-GUS construct
Plasmid pGALIPNIST
Plasmid pGALIPNIST
Plasmid pm16-1. C
Plasmid pm16. Chi
Vector pCX104 DNA
Vector pGN8 DNA se
Vector pGX52 DNA s
Vector pGX18 DNA s
Vector pGX22 DNA s
Plasmid PAN336 for
Chimeric gene cons
Plasmid PAN296 for
Plasmid PAN294 for
Haemophilia B gene
Retro virus vector
Retro virus vector
Intron 21 of human
Plasmid pB5/35Sbar
Human ovarian tumo
Plasmid pSEC-cyt/c
Vector pTGF67. S
Vector pTGF67. S
Human wild-type fa
Human factor VIII
Human factor VIII

XX
PS
XX

Disclosure; Column 87-90; 77pp; English.

This invention describes a novel energy transfer dye (I) which comprises a donor dye which is capable of absorbing light at a first wavelength and emitting excitation energy in response, and an acceptor dye which is capable of absorbing the excitation energy emitted by the donor and fluorescing at a second wavelength. (I) are used in processes for sequencing nucleic acids. (I) can be used in a wide variety of methods for detecting the presence of a component in a sample by labeling the component with (I). They are especially useful in processes which combine separation with fluorescent detection techniques, particularly methods requiring the simultaneous detection of multiple overlapping analytes. (I) have a strong fluorescent signal, the orientation between the donor and acceptor fluorophores is such that the energy is efficiently transferred between them. This sequence represents a fragment of plasmid pGEM DNA which is used to describe the method of the invention.

Sequence 738 BP; 170 A; 208 C; 192 G; 168 T; 0 other;

Query Match 100.0%; Score 738; DB 20; Length 738;
Best Local Similarity 100.0%; Pred. No. 1.6e-124;
Matches 738; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATACACTCATTAGGCGGAATTCAGCTCGGTACCGGGATCCTCTAGAGTCGACCT 60
Db 1 atacactcattagggcggaatttcagctcggtacccgggatacctctctagagtcgaact 60
QY 61 GCAGGCGATGCAAGCTTGTATCTATAGTTCACCTAAATAGCTGGCGTAACTCATGT 120
Db 61 gcaggcatgcaagctgtgtatctatagttcactaaatagctggcgtaactcatgt 120
QY 121 CATAGCTGTTCTGTGTAATGTTATCCGCTCACAAATCCACACATACAGCGCG 180
Db 121 catagctgttctgtgtatgtaatttcctgtcacaaatccacacatacagcg 180
QY 181 GAAGCATAAAGCTGTAAGCTGGGTGCTATGAGTGAAGTCACTCAATTAATGCT 240
Db 181 gaagcataaagctgtatgtaagctgggtgctatgagtgaggtgagctcaatttaattgct 240
QY 241 TCGCTCACTGCCGCTTTCAGTCCGGGAAACCTGCTGCGCAGCTGCTAATTAATGCT 300
Db 241 tgcgtcaactgccgcttccagtcgggaacctgtctgagctgcttaattgaatgc 300
QY 301 GCCAAGCGCGGGAGAGCGGTTTGCATTTGGCGCTTCTCCGCTTCTCCGCTCACTG 360
Db 301 gccaaagcgcgggagagcggtttgcatttggcgctctctccgcttctccgctcaactg 360
QY 361 ACTGCTCGCTCGGTCGCTGCTGCTGCGGAGCGGTATCAGCTCACTCAAGGCGGTAA 420
Db 361 actgctcgctcggtcgctgctgctgctgctgctgctgctgctgctgctgctgctgct 420
QY 421 TACGTTATCCACAGATCAGGGAATACGAGGAAAGAAACATGTGAGCAAGGCCAGC 480
Db 421 tacggttatccacagatcagggaatcagaggaagaaacatgtgagcaagggcgagc 480
QY 481 AAAAGGCCAGAACCGTAAAGCGCGGTTGCTGCGGCTTTTTCATAGGCTCCGCCGCC 540
Db 481 aaaagggccaggaacgtaaaagcgcggttgctgctgctgctgctgctgctgctgctgct 540
QY 541 CTGACGAGCATCACAAAATCGAGCTCAAGTCAGAGGTGGCGGAACCGACGACATAT 600
Db 541 ctgacgagcatcacaaaatcgagctcaagtcagaggtggcgaaacccgacagactat 600
QY 601 AAAGATACGAGCGGTTTCCCTTGGAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 aaagatacagagcggtttcccttggagctccctgctgctgctgctgctgctgctgctgct 660
QY 661 CGCTTACCGGATACCTGCTCCGCTTTTCTCCCTTGGGAGCGGTGGCGCTTCTCATAGCT 720
Db 661 cgcttacggatacctgctccgcttttctcccttgggagcggtggcgcttctctcatagct 720

QY 721 CACGCTGTAGTATCTCA 738
Db 721 cacgctgtagtatctca 738

RESULT 2

AAC83524
ID AAC83524 standard; DNA; 972 BP.

AC AAC83524;

DT 28-FEB-2001 (first entry)

XX T7 promoter sequence PCR primer SEQ ID NO: 3.

XX Diagnostic sequencing; specific cleavage; mass spectrometry;
KW disease diagnosis; polymorphism identification;
KW nucleic acid identification; PCR primer; ss.

OS Bacteriophage T7.

PN WO200066771-A2.

XX 09-NOV-2000.

XX 30-APR-2000; 2000WO-EP03904.

XX 30-APR-1999; 99US-0131984.

XX (METH-) METHEXIS NV.

XX Zabeau M, Stanssens P;

XX WPI: 2000-672835/65.

Sequence analysis of target nucleic acids for clinical and public health studies, such as diagnosing genetic diseases, comprises comparing mass spectra of the target nucleic acids with a known reference to detect sequence differences.

Example 2; Fig 5; 103pp; English.

The present invention relates to a novel method of sequencing nucleic acids involving the use of specific cleavage and mass spectrometry. This can be used particularly to re-sequence known nucleic acids when searching for polymorphisms or mutations. It can be used in disease diagnosis, in genotyping, to determine sequences responsible for drug resistance and pathogenicity, to identify and classify nucleic acids and in the determination of identity or heredity. In particular, it can be used in the diagnosis of haemophilias, thalassemias, Duchenne muscular dystrophy, Huntington's and Alzheimer's diseases and cystic fibrosis, and to diagnose predisposition to diabetes, atherosclerosis, obesity, autoimmune diseases and cancer.

Sequence 972 BP; 217 A; 272 C; 261 G; 222 T; 0 other;

Query Match 98.0%; Score 723; DB 21; Length 972;
Best Local Similarity 100.0%; Pred. No. 7.8e-122;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GGGCGAATTCGAGCTCGGTACCGGGGATCTCTAGAGTCGACCTGCAGCGATGCAAGCT 75
Db 1 gggcgaaatcagctcggtacccgggatacctctagagtcgacctgcagcgatgcaagct 60

QY 76 TCAGTATTTCTATAGTCTACCTAAATAGCTTGGCGTAATCATGCTCATAGCTGTTTCCTG 135
Db 61 tgagtattctatagcttcacctaaatagcttggcgtaacatgctgctgctgctgctgct 120

QY 136 TGTGAATTTGTTATCGCTCACAATTTCCACACATAGAGCGGGAAGCAATAAGTGTA 195
Db 121 tgtgaatttgttatcgctcacaaatccacacatacagcgccggaagcataaagtata 180

Qy 196 AAGCTGGGGTGCCTAATAGTGAAGTAACTACATTAATTTGGCTGCTACTGCCG 255
 Db 181 aagcctgggtgcttaagtgaagctaaactcaatcaattggtgctgctactgccc 240
 Qy 256 CTTTCCAGTGGGAAACCTGCGGCGAGCTGCAATTAATGAATCGGCAACGCGGGGA 315
 Db 241 cttccagtcgggaaacctgctgcccagctgcaattgaatcggccacgcgcgggga 300
 Qy 316 GAGCGGTTTGGCTATTGGGCGCTCTTCCGCTTCCCTGCTACTGACTCGCTGCGCTCG 375
 Db 301 gagcgggttgctgattggcgctctccgcttccctgcctcactgactgcgtcg 360
 Qy 376 TCGTTGGCTCGGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATCCAC 435
 Db 361 tcgttcggctgcgagcggtatcagctcactcaaaaggcgtaatacgttatccacg 420
 Qy 436 AATCAGGGGATAACGAGGAAGAACATGTGAGCAAAAGGCGCAAGAGGCAAGAAC 495
 Db 421 aatcaggggataacgcgggaaagaaacatgtagcaaaaggcgtaatacgttatccacg 480
 Qy 496 GTAAAGAGCGCGCTTGTGCGCTTTTTCATAGGCTCCGCCCCCTTGACGAGCATACA 555
 Db 481 gtaaaaggcgctgtgctggtttttccatagctccgccccctgagcagcatcaca 540
 Qy 556 AAATCGAGCTCAAGTCAGAGTGGCGGAACCCGACAGGACTATAAGATACCGGGT 615
 Db 541 aaatcgagctcaagtcagaggtggcgaaacccgacagggactataagataccaggcgt 600
 Qy 616 TTCCCGCTGGAAGCTCCCTGCTGCGCTCTCTGTTCCGACCCCTGCGGTTACCGGATAC 675
 Db 601 ttcccgctggagctccctgctgctctctgctctctctgctcagacctgaccttaccggatacc 660
 Qy 676 TGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATC 735
 Db 661 tgtcgcctttctcccttcgggaagcgtggcgctttctcatagctcaogctgtaggtatc 720
 Qy 736 TCA 738
 Db 721 tca 723

RESULT 3
 ID AAV43439 standard; DNA; 3253 BP.
 AC AAV43439;
 DT 29-OCT-1998 (first entry)
 XX DNA sequence of the vector pGEM3-2f(-).
 KW recA gene; extein; screening; antimicrobial activity; intein;
 KW protein-splicing element; growth protein; antimicrobial agent;
 KW prevention; splicing; identification; conditionally-splicing; ss.
 OS Synthetic.
 XX US5795731-A.
 XX 18-AUG-1998.
 XX 26-AUG-1996; 96US-0702902.
 XX 26-AUG-1996; 96US-0702902.
 XX (HEAL-) HEALTH RES INC.
 XX Belfort M;
 XX WPI; 1998-466664/40.
 XX Screening agents for antimicrobial activity - by monitoring affect

PT of splicing intein into reporter gene, also useful for studying
 PT intein function
 XX
 PS Example 1; Fig 6B; 64pp; English.
 XX
 CC The present sequence represents the DNA sequence of the vector
 CC pGEM3-2f(-), used in the course of the invention. The specification
 CC describes a method for screening agents for activity against a microbial
 CC pathogen that has an intein (i.e. a protein-splicing element occurring
 CC naturally as an in-frame protein fusion) in a gene that encodes a protein
 CC that facilitates growth. The method comprises preparing recombinant
 CC clones of an inducible expression vector containing an altered reporter
 CC gene, including a silent restriction site, and the intein, and detecting
 CC the presence of the extein products of the agent. Reduced production of
 CC the extein product indicates inhibition of the intein, i.e. that the
 CC agent has antimicrobial activity. The agents of the invention represent
 CC a new type of antimicrobial agent that prevents splicing out of the
 CC intein and thus formation of an active, essential protein. The same
 CC method is used to identify conditionally-splicing inteins and to study
 CC intein function.

Sequence 3253 BP; 793 A; 817 C; 805 G; 838 T; 0 other;
 Query Match 98.0%; Score 723; DB 19; Length 3253;
 Best Local Similarity 100.0%; Pred. No. 6.3e-122;
 Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 GGGCGAATTCGAGTTCGGTACCCGGGATCTCTAGAGTGCACCTGAGCATGCAAGCT 75
 Db 55 gggcgaaattcgagctggtaccggggtactctctagagtcgacgtcagcatgcaagct 114
 Qy 76 TGAGTATCTATAGTGCACCTAATAGTTCGGCTAATCATGTGCTATAGTGTTCCTG 135
 Db 115 tgagtattctatagtgcaacctaaatagcttggtgctaatcatggtcatagctgttctcg 174
 Qy 136 TGTGAAATTTGTTATCCGCTCACAAATTCACACATACGAGCGGGAAGCATAAAGTGA 195
 Db 175 tgtgaaattgttatccgctcacattccacacacacacagcgcggaagcataaagtga 234
 Qy 196 AAGCTGGGTGCTCTAATAGTGCATTAATTCGCTTGGCTGCTACTGCCG 255
 Db 235 aagcctgggtgctctaattagtgagctaaactcaattatgcttgcgtcactgctcgcg 294
 Qy 256 CTTTCCAGTGGGAAACCTGCTGCGAGTGCATTAATGAATCGGCAACGCGGGGA 315
 Db 295 cttccagtcgggaaacctgctgcccagctgcaattaatgaaatcgccaaacgcgcgggga 354
 Qy 316 GAGCGGTTTGGCTATTGGGCGCTCTTCCGCTTCTCTGCTCACTGACTGCTGCTGG 375
 Db 355 gagcgggttgctgattggcgctctctcctcctcctcactgactcgtcgtcgtcgcg 414
 Qy 376 TCGTTCGCTCGGCGGAGCGGTATCAGCTCACTCAAGCGGTATACGTTATCCACAG 435
 Db 415 tcgttcggctgcgagcggtatcagctcaactcaaaaggcggttaacggtttatccacag 474
 Qy 436 AATCAGGGGATAACGAGGAAGAACATGTGAGCAAAAGGCGCAAGAGGCAAGAAC 495
 Db 475 aatcaggggataaacgcgggaaagaaacatgtagcaaaaggcgcaaaaggccaggaacc 534
 Qy 496 GTAAAGAGCGCGTTCGCTGCGGTTTTCATAGGCTCCGCCCCCTTGACGAGCATACA 555
 Db 535 gtaaaaggcgcggtgctggtttttccatagctcgcgcctcgtcagcagcatcaca 594
 Qy 556 AAATCGAGCTCAAGTCAAGTGCAGAGTGGCGGAACCCGACAGGACTATAAGATACCGGCT 615
 Db 595 aaatcgagctcaagtcagaggtggcgaaacccgacagggactataaagataccagggcgt 654
 Qy 616 TTCCCGCTGGAAGCTCCCTGCTGCGCTCTCTCTGTTCCGACCCCTGCGGTTACCGGATAC 675
 Db 655 ttcccgctggagctccctcgtcgtcctctctctctctcgttccgacctcgcgttaccggatacc 714

QY 676 TGTCGGCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAGCGTGTAGGTATC 735
 Db 715 tgtcgccttctcccttcggaagcgtggcgttctctcatagctcagcgtgtaggatc 774
 QY 736 TCA 738
 Db 775 tca 777

RESULT 4
 AAZ49993
 ID AAZ49993 standard; DNA; 3216 BP.
 AC AAZ49993;
 XX
 XX
 DT 25-APR-2000 (first entry)
 XX
 DE Plasmid pGNI for expression of dsRNA in *C. elegans*.
 XX
 KW Double stranded RNA inhibition; dsRNA; transgenic organism; MCS;
 KW pest infestation; plasmid pGNI; T7 promoter; multiple cloning site; ds.
 XX
 OS Synthetic.
 OS Bacteriophage T7.
 PN WO200001846-A2.
 XX
 XX
 PD 13-JAN-2000.
 XX
 PF 02-JUL-1999; 99WO-EP04718.
 XX
 PR 03-JUL-1998; 98GB-0014536.
 PR 09-DEC-1998; 98GB-0027152.
 XX
 XX
 PA (DEVG-) DEVGEN NV.
 XX
 PI Plaetinck G, Platteeuw C, Mortier K, Bogaert T;
 XX
 DR WPI; 2000-171025/15.
 XX
 PT Characterizing and assigning gene function, and providing a mechanism
 PT for alleviating pest infestation of plants -
 XX
 PS Claim 49; Fig 1; 97pp; English.
 XX
 CC The patent discloses a method for characterisation and assigning gene
 CC function and identifying DNA responsible for inducing a specific
 CC phenotype, using double stranded RNA inhibition (dsRNAi), which results
 CC in potent and specific interference with expression of endogenous genes
 CC in a cell. The method can be used for producing transgenic organisms
 CC and provides a selective mechanism for alleviating pest infestation of
 CC plants, such that the expressed ds RNA in the plant inhibits the
 CC expression of the DNA within the pest which is critical for its growth,
 CC survival, proliferation or reproduction. The present sequence is plasmid
 CC pGNI, comprising two T7 promoters oriented towards each other with a
 CC multiple cloning site (MCS) between them. This plasmid can be used to
 CC transform *Caenorhabditis elegans* to induce expression of dsRNA for
 CC identifying DNA responsible for conferring a particular phenotype.
 XX
 SQ Sequence 3216 BP; 827 A; 788 C; 821 G; 780 T; 0 other;

Query Match 93.7%; Score 691.8; DB 21; Length 3216;
 Best Local Similarity 96.2%; Pred. No. 2.6e-116;
 Matches 726; Conservative 0; Mismatches 12; Indels 17; Gaps 1;

QY 1 ATACGACTCATATAGCGCGAATTCGAGCTCGGTACCGGGGATCCTCTAGAGTCGACCT 60
 Db 686 atagactcatatagggcgaaatcgagctcggtaccgcgggagtcctctagagtcgaaag 745
 QY 61 GCAGGC-----ATGCAAGCTTGAGTATCTATAGTCTCACCCTAAATAG 103
 Db 746 cttctgcctatagtgagtcgtattacagcttgagttattctatagtgatcacctaaatag 805

QY 104 CTTGGCGTAATCATGTCATAGCTGTTCTCTCTGTGAAATTTGTTATCCGCTCACAATTC 163
 Db 806 ctggcgtaatcatggtcatagctgttctctgtgaaattgttatccgctcacaaatcc 865
 QY 164 ACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGTGCTTAATGATGAGCTA 223
 Db 866 acacaacatagcgagcggaagcataaagtgtaaagcctggggtgcttaatgagtgagccta 925
 QY 224 ACTCACATTATTTGGCTTGGCTCAGTCCCGCTTTCAGTTCAGTTCGGGAAACCTGTCTG 283
 Db 926 actcacattaattgctgtgctcactgcccgtcttccagtcggaacactgctgctgcca 985
 QY 284 GCTGCATTAAATCAATCGGCCAACGCGGGGAGAGCGGTTTTCGTTATTTGGCGCTCTTC 343
 Db 986 gctgcattaaatgaatcgcccaacgcgagggagagcgttctgctattggcgctcttc 1045
 QY 344 CGTTCCTCGTCTACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 403
 Db 1046 cgcttctcgtcactgactgctgctgctgctgctgctgctgctgctgctgctgctgct 1105
 QY 404 TCACTCAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATACGAGGAAACAT 463
 Db 1106 tcaactcaagcggttaatacaggttatccacagaaatcaggggataaacgcaggaaagaacat 1165
 QY 464 GTGAGCAAAAGCCGCAAAAGCGGAGGAGCGGAAACCGTAAAGCGCGTTCGCTGGCGTTT 523
 Db 1166 gtgagcaaaagccagcaaaagccaggaacgcgttaaaagccgctgctgctgctgctgct 1225
 QY 524 CCATAGGCTCGGCGCCCTGAGCAGCATATCAAAATCGAGCTCAAGTCAAGAGTGGCG 583
 Db 1226 ccataggctcggcgccctgagcagcatcaaaaaatcgagcgtcaagtcagaggtgctg 1285
 QY 584 AACCCGACAGGACTATAAGATACAGGCGCTTCCCGCTCGAAGCTCCCTCGTGGCTC 643
 Db 1286 aaaccgacagggactataagataccagggcgttcccccctggaagctccctcgctgctc 1345
 QY 644 TCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCCCTTCTCTCCCTTCGGGAAGCGT 703
 Db 1346 tcctgttcgacccctcgcttaccggataacctgctgcttcccttccttccttcgggaagcgt 1405
 QY 704 GCGCTTCTCATAGCTCAGCTCAGCTAGTATCTCA 738
 Db 1406 ggcgcttctcatagctcagcgttaggtatctca 1440

RESULT 5
 AAS06382
 ID AAS06382 standard; DNA; 3216 BP.
 XX
 AC AAS06382;
 XX
 DT 23-OCT-2001 (first entry)
 XX
 DE Vector pGNI DNA sequence relating to double-stranded RNA inhibition.
 XX
 KW Vector pGNI; bacteriophage T7 promoter; multiple cloning site; MCS;
 KW ampicillin resistance; nematode worm; *Caenorhabditis elegans*; RNAi;
 XX double-stranded RNA inhibition; circular; cyclic; mutant; ds.
 OS Bacteriophage T7.
 OS Synthetic.
 XX
 PN WO200148183-A2.
 XX
 PD 05-JUL-2001.
 XX
 PF 22-DEC-2000; 2000WO-EP13149.
 XX
 PR 24-DEC-1999; 99GB-0030691.
 XX
 PA (DEVG-) DEVGEN NV.
 XX

Db 1286 aaaccgacagagactataaagataccagcgctttccccctggagctccctcgtgogctc 1345
OY 644 TCCTGTTCCGACCGCTGCCGCTTACCGGATACCTGTCCGCGCTTCTCCCTTCGGGAAGCGT 703
Db 1346 tctgttccgacacctgcgcttaccgagatacctgtccgcttctcccttcggaagcgt 1405
OY 704 GCGGCTTCTCATAGCTACCGCTGTAGGTATCTCA 738
Db 1406 ggcgctttccatagctccacgctgtaggtatctca 1440

RESULT 6

AAZ29606
ID AAZ29606 standard; DNA; 8126 BP.
XX
AC AAZ29606;
XX
DT 30-MAR-2000 (first entry)
XX
DE Basic fragmentation vector, pDVO DNA.
XX
XX Fragmentation vector; pDVO DNA; target sequence; triplet repeat; yeast;
XX lysine-2 gene; selectable marker; Yeast Artificial Chromosome; YAC;
KW telomere; homologous recombination; deletion fragment; sequencing;
KW End Rescue Site; ERS; positional information; config mapping; ds.
XX
OS Saccharomyces cerevisiae.
OS Synthetic.

XX Key Location/Qualifiers
FH misc_feature 1..14
FT /tag= a
FT /label= End_Rescue_Site
FT /note= "Contains restriction sites for EcoRI, BamHI,
FT KpnI and ClaI"
FT 15..4819
FT /tag= b
FT /label= Yeast_Lysine-2_marker
FT /function= "Selection marker"
FT /note= "Used for acentromeric fragmentation"
FT 4820..4961
FT /tag= c
FT /label= Yeast_telomere_sequence
FT /note= "Used for acentromeric fragmentation"
FT 4962..8126
FT /tag= d
FT /label= pGem3zf(-)_vector_sequence

XX WO9966059-A1.

XX 23-DEC-1999.

XX 11-JUN-1999; 99WO-EP04106.

XX 12-JUN-1998; 98EP-0201976.

XX (VLA-) VLAMS INTERUNIVERSITAIR INST BIOTECHNOG.

XX Del-Favero J, Van Broeckhoven C;

XX WPI; 2000-116773/10.

XX Fragmentation vector for producing deletion fragments of yeast
XX artificial chromosomes for genome mapping -
XX Claim 1; Fig 8; 52pp; English.

XX The present sequence is the basic fragmentation vector DNA, pDVO.
XX The vector comprises a yeast telomere sequence, a selectable marker,
XX yeast Lysine-2 gene and a repetitive element comprising triplet repeat
XX regions, allowing for homologous recombination, between the vector and
XX the YAC. It has SalI restriction site (4961-4966) for linearisation of
XX the vector, PstI site (4967-4972) for cloning of target sequence and

PI Plaetinck G, Mortier K, Lissens A, Bogaert T;
XX WPI; 2001-425660/45.
XX Inhibiting target gene expression in nematode worm by feeding it a food
XX organism capable of producing double-stranded RNA structure having a
XX sequence identical to target gene following ingestion of the food -
XX Example 1; Page 50-51; 60pp; English.

XX The present sequence for vector pCNI comprises opposable
XX bacteriophage T7 promoters flanking a multiple cloning site and
XX an ampicillin resistance marker. The sequence is described in an
XX invention relating to a method of inhibiting expression of a target
XX gene in a nematode worm (e.g. Caenorhabditis elegans). The method
XX involves feeding to the worm a food organism which is capable of
XX producing a double-stranded RNA structure having a nucleotide sequence
XX substantially identical to a portion of the target gene following
XX ingestion of the food organism by the nematode, or by introducing a DNA
XX capable of producing the double-stranded RNA structure. The nematode
XX worm has a non wild-type genetic background selected to provide
XX increased sensitivity to RNA interference as compared to wild type,
XX and exhibits increased gut uptake compared to wild type.

XX Sequence 3216 BP; 827 A; 788 C; 821 G; 780 T; 0 Other;

Query Match 93.7%; Score 691.8; DB 22; Length 3216;
Best Local Similarity 96.2%; Pred. No. 2.6e-116;
Matches 726; Conservative 0; Mismatches 12; Indels 17; Gaps 1;

OY 1 ATACGACTCATATAGGGCGAATTCGAGTCGGTACCCGGGATCTCTAGAGTCGACCT 60
Db 686 atacgactcatatagggcgaaatcgcagtcggtacccgggagctctctagatgcgaag 745
OY 61 GCAGGC-----ATCGAAGCTTGAGTATTCTATAGTGTCACTAATAATAG 103
Db 746 cttctgcctcatagtagtcgtattacagcttgatattctatagtgccacctaataag 805
OY 104 CTTGGCGTATATGTTGATAGTGTCTCTGTGTAATTTGTTATCCGCTCAATTC 163
Db 806 ctggcgtaatacgtcatagctgtttctgtgtgaattgttatccgctcacaattcc 865
OY 164 ACACACATACGAGCGGAGCATAAAGCTAAAGCTGGGGTCCCTAATGAGTGAGCTA 223
Db 866 acacacatacagcgagcggaagcataagtgtaaaagcctg999cgcccaatgagtgagcta 925
OY 224 ACTCACATTAATTGCGTTCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTC 283
Db 926 actcacattaatgtcgttcgctcactgcgcgtttccagtcg99aaacctgctgcga 985
OY 284 GCTGCATTAATGAATCGGCAACCGCGGGGAGAGGGGTTTGGTATTGGCGCTTTC 343
Db 986 gctgcattaatgaatcgcccaacgcgcg99gagag99ggtttgctgattg99cgctctc 1045
OY 344 CGCTTCCTCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 403
Db 1046 cgtctctcctcactgactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1105
OY 404 TCACCTAAAGCGGTATACGGTTATCCACAGAAATCAGGGGATTAACGCAGGAAGACAT 463
Db 1106 tcaactaaagcggtaataacggttatccacagaatcgggatacaacgcaggaagaacat 1165
OY 464 GTGAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGT 523
Db 1166 gtgagcaaaagccagcaaaagccagcaaaagccagcaaaagccagcaaaagccagct 1225
OY 524 CCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACCTCAAGTCAGAGTGCGG 583
Db 1226 cgataggctcgcgcgcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1285
OY 584 AAACCCGACAGACTATTAAGATACCAGCGGTTTCCCTCGAAGCTCCCTCGTGGC 643
Db 1285 aaacccgacagactatataagataccagcggtttccctcgaaagctccctcggtggcgcgc 1343

CC Sphi site (4973-4978) for isolation of cloned fragmented end in
 CC combination with the restriction site used to generate the plasmid.
 CC The fragmentation vectors can be used for producing deletion fragments
 CC of YACs. The vectors can be used for the sequencing of at least a part
 CC of the human genomic sequences present in the YACs. New markers from
 CC the chromosomes of the human genome can be isolated and positional
 CC information of isolated sequences can be provided, for use in contig
 CC mapping using these vectors.
 XX
 SO Sequence 8126 BP; 2268 A; 1705 C; 1858 G; 2295 T; 0 other;

Query Match 93.4%; Score 689.4; DB 21; Length 8126;
 Best Local Similarity 99.9%; Pred. No. 6e-116;
 Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 48 CTAGAGTCGACCTGCGGATGCAAGCTTGTATATATAGTCTACCTAAATAGCTTG 107
 Db 4956 cttagagtcgaacctgacgcatgcaagcttgatattctatagtgctacacataatagcttg 5015

QY 108 GCGTAATCATGGTTCATAGCTGTTTCCCTGTGTGAATTTGTTATCCGCTCACAAATTCACAC 167
 Db 5016 gcgtaatcatggtcatagctgtttccctgtgtgaattgttatccgctcaccaattccacac 5075

QY 168 AACATAGCGGCGGAGCAATAAGTGTAAAGCCCTGGGCTGCCTAATGAGTGAGCTAACTC 227
 Db 5076 aacatagcgcggaagcataaagtgtaaagccctgggctgcctaatgagtgagctaacctc 5135

QY 228 ACATTAAATTCGGTTCGCTCCTACCTGCCGCTTCCAGTCGGGAAACCTGTGCGCAGCTG 287
 Db 5136 acattaatgcttgcgtcctacgctccgcttccagtcgggaaacctgtgctgagctcgt 5195

QY 288 CATTAAATGATCGCCCAACGCGGGGAGAGCGGCTTTCGTTGCTATGGCGCTCTTCGCGCT 347
 Db 5196 cattaatgattcgcccaacgcgggagagcggttgcgttgcgtatggcgctcttcgctcgt 5255

QY 348 TCCTCGCTCAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 407
 Db 5256 tctcgtcctcagctactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 5315

QY 408 TCAAGGCGGCTAATACGGTTCATCCAGATATCAGGGGATACAGGATACAGGATACAGGATAC 467
 Db 5316 tcaagcgcggttaatacggcttatccacagatcaggggataacgcaggaagaacatgtga 5375

QY 468 GCAAAAGGCCACCAAAAGCCAGGACCTGTAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 527
 Db 5376 gcaaaaggccacaaagccaggaacgctgaaagcgcgctgctgctgctgctgctgctgctgct 5435

QY 528 AGGCTCCGCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGGAAAC 587
 Db 5436 aggctccgccccctgacgagcatcacaaaaatcagctcaagtcagagtggtggcgaaac 5495

QY 588 CCGACAGGACTATAAGATACAGCGTTTCCTCCCTGGAAGCTCCCTGCTGCTGCTGCTGCTGCT 647
 Db 5496 ccgacaggaactataagataaccagcgcttccctccctggaagctccctgctgctgctgctgct 5555

QY 648 GTTCGAGCCCTGCGCTTACCGGATACCTGCTCCGCTTCCCTTCGCGGAGCGCTGCGG 707
 Db 5556 gtccgacctgctgcttaccggaataccctgctccgcttctcccttcggtggaagcggtggcg 5615

QY 708 CTTTCTCATAGCTACGCTGTAGTATCTCA 738
 Db 5616 ctttctcatagctacgctgtaggtatctca 5646

RESULT 7
 AAT86449
 ID AAT86449 standard; DNA; 4283 BP.
 AC AAT86449;
 XX
 DT 10-FEB-1998 (first entry)
 XX

DE DNA encoding hGH, EGF fusion protein from pWRG1630.
 XX Mature epidermal growth factor; EGF; wound healing; skin; gold bead;
 KW epidermal cell; hair follicle; polyA addition sequence;
 KW human growth hormone secretory signal peptide; hGH; ds.
 XX
 OS Chimeric - homo sapiens.
 OS Chimeric - bos taurus.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 713..1253
 FT /tag= a
 FT /note= "contains intronic sequences"
 FT sig_peptide 713..1049
 FT /tag= b
 FT /standard_name= secretion_signal
 FT /note= "human growth hormone secretory signal contains
 intronic sequences"
 FT mat_peptide 1050..1250
 FT /tag= c
 FT /product= mature_human_EGF
 FT 713..723
 FT /tag= d
 FT /number= 1
 FT intron 724..982
 FT /tag= e
 FT /number= 1
 FT exon 983..1253
 FT /tag= f
 FT /number= 2
 FT /note= "end of exon is given as TAG stop codon"
 XX
 PN US5661132-A.
 XX
 PD 26-AUG-1997.
 XX
 PF 22-NOV-1994; 94US-0343401.
 XX
 PR 14-DEC-1989; 89US-0451957.
 PR 22-MAY-1991; 91US-0707248.
 PR 11-JUN-1992; 92US-0897357.
 PR 11-JUN-1993; 93US-0076550.
 XX
 PA (AURA-) AURAGEN INC.
 XX
 PI Andree C, Eriksson E, Macklin MD, Swain WF;
 XX
 DR WPI; 1997-434422/40.
 DR P-PSDB; AAW32221.
 XX
 PT Genetic construct encoding mature epidermal growth factor - when
 PT delivered to cells on gold beads is useful for treatment of skin
 PT wounds
 XX
 PS Claim 3; Columns 11-18; 15pp; English.
 XX
 CC This sequence represents a new genetic construct consisting of, in
 CC 5'-to-3' order: a promoter effective in skin cells; a signal sequence
 CC that directs secretion of a protein from a skin cell; a non-naturally
 CC occurring DNA fragment which encodes an epidermal growth factor (EGF) and
 CC lacks all coding regions of a natural EGF precursor gene except the
 CC coding region that encodes the mature EGF; and a polyA addition sequence.
 CC The genetic construct is used in a method for treating a skin wound site
 CC on an animal. The presence of the EGF accelerates, by a statistically
 CC significant amount, the healing time of a wound treated in this manner.
 CC Also the need to use expensive purified mature EGF protein is avoided.
 XX
 SQ Sequence 4283 BP; 1023 A; 1095 C; 1114 G; 1051 T; 0 other;

Query Match 91.1%; Score 672.4; DB 18; Length 4283;
 Best Local Similarity 99.9%; Pred. No. 7.6e-113;


```

QY 545 CGAGCATCACAAAATCGACGCTCAAGTCAGAGTGCGGAAACCCGACAGACTATATAAG 604
Db 2070 cgagcatcacaaaatcgacgctcaagtcagagtggtgcgaacccgacagactataag 2129
QY 605 ATACCAGGGTTCCCTCGAAGCTCCCTCGTGGCTCTCTCTCTCTCTCTCTCTCTCTCT 664
Db 2130 ataccagggttccctcgaaagctccctcgtggtctctctctctctctctctctctct 2189
QY 665 TACCGGATACCTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 724
Db 2190 taccggatacctgctgctctctctctctctctctctctctctctctctctctctctc 2249
QY 725 CTGTAGGTATCTCA 738
Db 2250 ctgtaggtatctca 2263

RESULT 9
AAA63237
ID AAA63237 standard; DNA; 4283 BP.
AC AAA63237;
XX
XX
DT 21-NOV-2000 (first entry)
DE Plasmid pWRG1630 nucleotide sequence.
XX
KW Plasmid pWRG1630; epidermal growth factor; EGF; human; cytomegalovirus;
KW growth hormone; hGH; bovine; gene delivery; circular; microseeding; ds.
OS Chimeric - Homo sapiens.
OS Chimeric - Bos taurus.
OS Chimeric - Cytomegalovirus.
XX
XX
Key Location/Qualifiers
CDS 713..1253
FT /*tag= a
FT /product= "hGH/EGF fusion protein"
FT 713..721
FT /*tag= b
FT /number= 1
FT sig_peptide 713..1049
FT /*tag= c
FT Intron 722..980
FT /*tag= d
FT /cons_splice= (5'site:NO,3'site:NO)
FT 981..1250
FT /*tag= e
FT /*tag= f
FT mat_peptide 1050..1250
FT /*tag= f
FT /product= "hGH and plasmid polylinker and human EGF"
XX
XX
US06090790-A.
XX
XX
PD 18-JUL-2000.
XX
XX
PF 15-DEC-1997; 97US-0990442.
XX
XX
PR 14-DEC-1989; 89US-0451957.
PR 22-MAY-1991; 91US-0707248.
PR 11-JUN-1992; 92US-0897357.
PR 11-JUN-1993; 93US-0076550.
PR 19-MAY-1995; 95US-0445265.
XX
XX
(ERIK/) ERIKSSON E.
XX
XX
PI Eriksson E;
XX
XX
DR WPI: 2000-531428/48.
DR P-PSDB; AAB13435.

```

XX

PT

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Query Match

Best Local Similarity

Matches 673; Conservative

Mismatches 1; Indels 0; Gaps 0;

QY 65 GCATGCAAGCTTGAGTATTCCTATAGTGTACCTAAATAGTGTGGCGTAATCATGGTCA 124

Db 1590 gcatgcaagcttgagtattctatagtcacctaataatagttggcgaatcattggtc 1649

QY 125 GCTGTTTCTGTGTAATTTGTTATCCGCTCAACAATCCACACAAATAGAGCCCGGAAG 184

Db 1650 gctgttctcgtgtgaattgtatccgtccacaaattccacacacacacacacacac 1709

QY 185 CATAAAGTGTAAAGCTGGGTGCTCTAAATAGTGTAGTGTAACTCAATTAATTCGCTGG 244

Db 1710 cataaagtgtaaagcctgggtgcttaagtgtgagtaactcaatgaattgctgctg 1769

QY 245 CTCACCTGCCCGCTTTCCAGTCGGGAAACCTGTCGTCAGCTGCGTGCATTAATGAATCGGCC 304

Db 1770 ctcaactgcccgctttccagtcgggaaacctgctgtgagtcgagtcgtaataatgaatgcg 1829

QY 305 ACGCGCGGGAGAGCGGTTTGCCTATTGGGCGCTTCCTCCGCTTCCTCGCTCACTGACTC 364

Db 1830 acgcgcgggagagcggttgcgtatggcgctcttcctcgtctcctcgtcactgactc 1889

QY 365 GCTGCGCTCGGTCGTTTCGCTGCGGCGAGCGGTATCAGCTCACTCAAAAGCGGGTAATACG 424

Db 1890 gctgctcgtcgttcgttcggtgcgagcggtatcagtcactcaaaaggcggtlaatac 1949

QY 425 GTTATCCACAGATCAGGGGATAACGCGAGGAAAGAACATGTGACCAAAAGGCCAGCAAAA 484

Db 1950 gttatccacagaatcaggggataacgcaggaagaacatgagcagaagaagccagcaaa 2009

QY 485 GCCCAGGAACCGTAAAGAGCCCGCTTGTGCGCTTTTTCATAGGCTCGCCCGCCCTGA 544

Db 2010 gccagggaacccgttaaaaggccggttgcgtggttcttcgtaggtcgcgcgcctcga 2069

QY 545 CGAGCATCACAAAATCGAGCTCAAGTCAGAGTGCGGAAACCCGACAGACTATATAAG 604

Db 2070 cgagcatcacaaaatcgagctcaagtcagagtggtgcgaacccgacaggaactataag 2129

QY 605 ATACCAGGGCTTTCCCGCTCGAAGCTCCCTCGTGGCTCTCTCTCTCTCTCTCTCTCTCT 664

Db 2130 ataccaggcgtttcccccctggaagctccctcgtgcgtctcctcgttccgacctcgcgt 2189

QY 665 TACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTTCATAGCTCAGG 724

Db 2190 taccggatacctgctcgcctttctccttcgggaagcgtcggcgcttctcattagctcacg 2249

QY 725 CTGTAGGTATCTCA 738

XXXXXXXXXXXX

Delivering expressible genetic material into a perosteal cell useful for stimulating bone growth comprises repeatedly injecting the genetic material into the tissue at an in situ site on a human or non-human animal with a microneedle.

Example 2; Columns 17-22; 2lpp; English.

The present sequence is the nucleotide sequence of plasmid pWRG1630. This plasmid is an epidermal growth factor (EGF) encoding expression plasmid, and contains an in-frame fusion of the human growth hormone (hGH) secretory signal peptide to the mature human EGF peptide. The present sequence also includes a promoter from cytomegalovirus. Downstream of the mature human EGF coding region is the 3' untranslated sequence and polyadenylation site from bovine GH gene. Plasmid pWRG1630 was used in the direct gene transfer of EGF into skin target sites using a microneedle delivery apparatus: "microseeding". The microseeding method of the present invention is useful for treating various disorders such as burns, pain, tumours, infections and skin diseases.

Sequence 4283 BP; 1023 A; 1095 C; 1114 G; 1051 T; 0 other;

Db 2250 ctgtaggtatctca 2263

RESULT 10

AAAT69188

ID AAAT69188 standard; DNA; 4118 BP.

XX AAAT69188;

XX 26-FEB-1998 (first entry)

XX Construct pGEM-hTR containing RNA component of human telomerase.

XX Human telomerase; quantification; tumour cell; pGEM-hTR;

KW detection; micrometastasis; diagnosis; lymphoblastoma; leukaemia;

KW teratocarcinoma; melanoma; carcinoma; cancer; tumour; neuroblastoma;

KW rhabdomyosarcoma; leiomyosarcoma; lymphoma; RNA component; ss.

XX OS Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

FT misc_feature 12..975

FT /*tag= a

FT /note= "cDNA from RNA component of human telomerase"

XX W09718322-A2.

PN 22-MAY-1997.

XX 14-NOV-1996; 96WO-DE02183.

XX 16-NOV-1995; 95DE-4042795.

XX (DAHM/) DAHM M W.

XX Dahm MW;

XX WPI; 1997-289298/26.

XX Quantifying tumour cells in body fluid - by measuring RNA component

PT of telomerase after amplification, especially useful for early

PT diagnosis of metastasis

XX Example 5; Fig 5; 46pp; German.

XX The present sequence is the construct pGEM-hTR, which comprises the

CC transcription vector pGEM-132f(+) and the cDNA from the RNA

CC component of human telomerase. The construct was used in the

CC development of a novel method for quantifying tumour cells in a

CC body fluid. The method comprises specific amplification of the RNA

CC component of telomerase, and measuring the amount of amplified

CC nucleic acid. At least 1, preferably all 3 (AAAT69173-75) standard

CC nucleic acids are present (at different concentrations) and are

CC co-amplified with telomerase DNA. The amplification products are

CC detected directly or via a label or by hybridisation with a

CC labelled oligonucleotide (AAAT69184-87, 1 for each standard and 1 for

CC telomerase), and the amount of telomerase product compared with

CC that from the standards.

CC The method can be used to detect tumour cells, specifically

CC micrometastases, in a body fluid, particularly for the early

CC diagnosis of metastatic spread and for monitoring tumour therapy.

CC Typical tumour cells that can be quantified are micrometastases,

CC T cell lymphoblastoma or leukaemia (chronic myelogenous, and

CC chronic or acute lymphatic leukaemia), teratocarcinoma, melanoma,

CC carcinoma of lung, liver or prostate, cancer of colon or breast,

CC kidney, brain or adrenal tumours, neuroblastoma, rhabdomyosarcoma,

CC leiomyosarcoma and/or lymphoma.

XX Sequence 4118 BP; 991 A; 1032 C; 1094 G; 1001 T; 0 other;

Query Match 90.7%; Score 669; DB 18; Length 4118;

Best Local Similarity 97.1%; Pred. No. 3.1e-112;

Matches 681; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 38 CGGGATCTCTAGATGTCAGCATGCAAGCTTCTAGTATCTTATAGTGTACCT 97

Db 942 cgatgacctgagcagtaggataataacccccacaagcttgatattctatagtgca 1001

QY 98 AATAGCTTGGCGTAATCATGTCATAGCTGTTTCCCTGTGTGAATTTGTTATCGCTCAC 157

Db 1002 aatagcttgagcagtaggataataacccccacaagcttgatattctatagtgca 1061

QY 158 AATTCACACAAATACGAGCGGAGCAATAAAGTGTAAAGCTTGGGTGCTTAAATGAGT 217

Db 1062 aattccacacacatacagcgcgagcagcagcagcagcagcagcagcagcagcagc 1121

QY 218 GAGTAATCACTCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 277

Db 1122 gagctaacacataatgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1181

QY 278 GTGCAGCTGCATTAATGATCGGCAACGCGCGGAGAGCGGCTTTCGCTATTGGGCG 337

Db 1182 gtgcccagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1241

QY 338 CTCTTCCGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 397

Db 1242 ctcttcgcttctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1301

QY 398 ATCAGCTCACTCAAAAGCGGTAAATAGCTGTTATCCACAGCAATACGAGGATAACGAGGAAA 457

Db 1302 atcagctcaactcaaaagcggtgctgctgctgctgctgctgctgctgctgctgctgct 1361

QY 458 GAACATGTGAGCAAAAGCGGAGCAAAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 517

Db 1362 gaacatgtgagcaaaagcggtgctgctgctgctgctgctgctgctgctgctgctgct 1421

QY 518 GTTTTTCATAGGCTCCGCGCGCTTACCGGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 577

Db 1422 gtttttcataagctccgccccctgacgagcagcagcagcagcagcagcagcagcagcag 1481

QY 578 GTGGGAAACCGACAGGACTATAAAGATACGAGGCTTTCCTCCCTGGAAGCTCCCTCGT 637

Db 1482 gtggcgaaacccgacagagactataaagataccagcggtttccctccctggaagctccctcgt 1541

QY 638 GCGCTCTCTGTTCCGAGCCCTGCCGCTTACCGGATACCTGCTGCTGCTGCTGCTGCTGCTGCT 697

Db 1542 gcgctcctctgttccgacctgacctgacctgacctgacctgacctgacctgacctgacct 1601

QY 698 AAGCTGGCGCTTCTCATAGCTCAGGCTGCTAGGTATCTCA 738

Db 1602 aagcgtggcgctttctcatagctcagcgtgtaggtatctca 1642

RESULT 11

AAAT69189

ID AAAT69189 standard; DNA; 4118 BP.

XX AAAT69189;

XX 26-FEB-1998 (first entry)

XX Construct pGEM-hTR(Ka) containing RNA component of human telomerase.

XX Human telomerase; quantification; tumour cell; pGEM-hTR(Ka);

KW detection; micrometastasis; diagnosis; lymphoblastoma; leukaemia;

KW teratocarcinoma; melanoma; carcinoma; cancer; tumour; neuroblastoma;

KW rhabdomyosarcoma; leiomyosarcoma; lymphoma; RNA component; ss.

XX OS Homo sapiens.

OS Synthetic.

XX W09718322-A2.

XX 22-MAY-1997.

XX 14-NOV-1996; 96WO-DE02183.
 XX 16-NOV-1995; 95DE-4042795.
 XX (DAHM/) DAHM M W.
 XX Dahm MW;
 XX WPI; 1997-289298/26.
 XX Quantifying tumour cells in body fluid - by measuring RNA component
 PT of telomerase after amplification, especially useful for early
 PT diagnosis of metastasis
 XX Example 5; Fig 6; 46pp; German.
 XX The present sequence is the construct pGEM-hTR(Ka), which comprises
 CC the transcription vector pGEM-13zf(+) and the cDNA from the RNA
 CC component of human telomerase. The construct was used in the
 CC development of a novel method for quantifying tumour cells in a
 CC body fluid. The method comprises specific amplification of the RNA
 CC component of telomerase, and measuring the amount of amplified
 CC nucleic acid. At least 1, preferably all 3 (AAT69173-75) standard
 CC nucleic acids are present (at different concentrations) and are
 CC co-amplified with telomerase DNA. The amplification products are
 CC detected directly or via a label or by hybridisation with a
 CC labelled oligonucleotide (AAT69184-87, 1 for each standard and 1 for
 CC telomerase), and the amount of telomerase product compared with
 CC that from the standards.
 CC The method can be used to detect tumour cells, specifically
 CC micrometastases, in a body fluid, particularly for the early
 CC diagnosis of metastatic spread and for monitoring tumour therapy.
 CC Typical tumour cells that can be quantified are micrometastases,
 CC T cell lymphoblastoma or leukaemia (chronic myelogenous, and
 CC chronic or acute lymphatic leukaemia), teratocarcinoma, melanoma,
 CC carcinoma of lung, liver or prostate, cancer of colon or breast,
 CC kidney, brain or adrenal tumours, neuroblastoma, rhabdomyosarcoma,
 CC leiomyosarcoma and/or lymphoma.
 XX SQ Sequence 4118 BP; 988 A; 1031 C; 1093 G; 1006 T; 0 other;

Query Match 90.7%; Score 669; DB 18; Length 4118;
 Best Local Similarity 97.1%; Pred. No. 3.1e-112;
 Matches 681; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
 QY 38 CGGGGATCCTCTAGAGTCGACCTGCAGGCAATGCAAGCTTGAGTATTCTATAGTGTACCT 97
 Db 942 cgaatgacattgagcagtaggagataaaccacacagcttgagttatctatagtgacact 1001
 QY 98 AATATAGCTTGGCGTAATCATGGTGCATAGCTGTTTCTCTGTGGAATGTTATCGGTCAC 157
 Db 1002 aaatagcttgagcgaataca199tcagctgcttccctgctgaaattgttatccgctcac 1061
 QY 158 AATTCACACACATACAGAGCGGAGCAATAAAGTGTAAAGCTTGGCTGCTAATGAGT 217
 Db 1062 aattccacacacacacagcagcgaacataaagtgtaaacgctgaggtgcttaagt 1121
 QY 218 GAGTACTACATTAATTCGGTTGCGCTCACTGCGCGCTTTCAGTTCGGGAAACCTGTC 277
 Db 1122 gagctactcacattatgcttgctgctgctgctgctgctgctgctgctgctgctgctgct 1181
 QY 278 GTCCAGCTGCATTAATGAATCGGCCAACGCGGGGAGAGGGTTCGTTATGGGCG 337
 Db 1182 gtccagctgcattaatgaatcgccacacgcg999gag999gctgctgctgctgctgctgct 1241
 QY 338 CTCTTCGGTTCCTCGTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 397
 Db 1242 ctcttcgcttctcgtcactgactgctgctgctgctgctgctgctgctgctgctgctgct 1301
 QY 398 ATCAGCTCACTAAAGCGGTATACGGTATATCCAGAAATCAGGGGATTAACGAGGAAA 457
 Db 1301 atcagctcactaaagcggtatagcgttatccagaaatcaggggatcaacgagggaaa 457

Db 1302 atcagctcactcaaaagcggtlaatacaggttatccacagaatcaggggataacgagga 1361
 QY 458 GAACATGTGAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAG 517
 Db 1362 gaacatgtgagcaaaagcggtlaatacaggttatccacagaatcaggggataacgagga 1421
 QY 518 GTTTTTCATAGGCTCGCCCTCGCTGACGAGCATCAAAAATCGAGCTCAAGCTCAGAG 577
 Db 1422 gtatttcataagctcgccctcgccctcgccctcgccctcgccctcgccctcgccctcg 1481
 QY 578 GTGGGAAACCCGACAGGACTATAAGATACACGAGCGCTTTCCCGCTGGAAGCTCCCTCGT 637
 Db 1482 gtggcgaacccgacagagactataaagataccagcgctttcccccctgaaagctccctcg 1541
 QY 638 GCGCTCTCTGTTCCGACCCCTGCGCTTACGGGATACCTGTCCGCGCTTTCCTTCGGG 697
 Db 1542 ggcctctctctgttcgacccctgcttaccgatacctgctgctgctgctgctgctgctg 1601
 QY 698 AAGCGTGGGCGCTTCTCTCATAGCTCAGCTGCTAGGTATCTCA 738
 Db 1602 aagcgtggcgctttctctcatagctcagcgtgtaggtatctca 1642
 RESULT 12
 AAQ87347
 ID AAQ87347 standard; DNA: 4539 BP.
 XX AAQ87347;
 XX AC AAQ87347;
 XX DT 19-SEP-1995 (first entry)
 XX DE Plasmid pINVI.
 XX KW Plasmid pINVI; intron: recombination; combinatorial gene;
 KW trans-splicing; (IVS5,6)-exon-(IVS1-3); gene therapy;
 KW tissue plasminogen activator; tPA; thrombolytic; ds.
 XX OS Synthetic.
 XX FH Key
 FT exon Location/Qualifiers
 FT 969..1259
 FT /*tag= a
 FT /*label= E3-exon
 FT 1290..1559
 FT /*tag= b
 FT /*label= Exon-E5
 XX PN W09507351-A.
 XX PD 16-MAR-1995.
 XX PF 12-SEP-1994; 94WO-US10146.
 XX PR 10-SEP-1993; 93US-0119512.
 XX (HARD) HARVARD COLLEGE.
 XX PI Jarrell KA;
 XX WPI; 1995-123425/16.
 XX New intron-mediated recombinant techniques - used for the
 XX generation and selection of novel genes and gene prods. for use
 XX in therapy
 XX Example 1; Page 53-56; 87pp; English.
 XX An intron fragment construct, (IVS5.6)-exon-(IVS1-3), can mediate
 XX trans-splicing between heterologous exons and the generation of
 XX circular RNA transcripts. It was synthesized from plasmid pINVI
 XX (given in AAQ87347). Its intronic sequences correspond to half
 XX molecules generated by interruption of the 5' intron of the yeast
 XX mitochondrial ox13 gene in domain IV; exonic sequences E5 and E3 are

CC naturally disposed at the 5' and 3' ends of the 5g intron,
 CC respectively. pINVI was used to engineer human tPA to improve
 CC thrombolytic activity.

XX Sequence 4539 BP; 1246 A; 979 C; 998 G; 1316 T; 0 other;

Query Match 90.5%; Score 667.8; DB 16; Length 4539;
 Best Local Similarity 97.9%; Pred. No. 5.1e-112;
 Matches 688; Conservative 0; Mismatches 12; Indels 3; Gaps 1;

QY 39 GGGGATCCTCTAGCTCGACCTGCGAGCATGCAAGCTTGAGTATTCTATAGTGTACACCA 98
 DB 2235 ggggattcctctagagtcgacctgcagcatgcaagctttgtcccttttagtgagggtta 2294
 QY 99 AAT---AGCTTGGCTTAATCATGTCATAGCTGTTTCTGTTGTAATTCCTGCTC 155
 DB 2295 atttcgagcttggtgtaatacattggtcattagctgtttcctgtggaattgttaccgctc 2354
 QY 156 ACAATTCCACACATAGCAGCGGAGCATATAAGTGTAAAGCCTGGGGTCCCTAATGA 215
 DB 2355 acaattccacacacatacagccggaagcataaagtgtaagcctgggggacctcaatga 2414
 QY 216 GTGAGTAACTCACATTAATTCGTTGCTCCTCACTGCCGCTTCCAGTCGGGAAACCTG 275
 DB 2415 gtgagctaaactacattaatgcttggtgctcactgcccgtttccagtcgggaaacctg 2474
 QY 276 TGTGTCACGCTGATTAAATGATGCGCCAAACGCGGGGAGAGCGGTTTGGCTATTGGG 335
 DB 2475 tcgtgcagctgcattaaatgaatcgccaaacgcggggagagcggtttgctgattggg 2534
 QY 336 CGCTCTTCGCTTCCTCGCTCACTGACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 395
 DB 2535 cgctcttcgcttcctcgctcactgactgctgctgctgctgctgctgctgctgctgctg 2594
 QY 396 GTATGAGTCACTCAAGCGGTTAATACGTTTATCCACAGATCAAGGAGTAAACGAGGA 455
 DB 2595 gtatcagctcactcaagcggttaatacgtttatccacagaaatcaggggataacgagga 2654
 QY 456 AGAAGATGTGACAAAGGCGGAGCAAGGCGGAGCAAGCGGTAAAGAGCCCGCTGCTG 515
 DB 2655 aagaacatgtgagcaaaaggcagcaaaaggcagcaaaaggcagcaaaaggcagctgctg 2714
 QY 516 GCGTTTTCATAGGCTCGCCCTGAGCAGCATCAAAAATCGAGCTCAAGTCAAG 575
 DB 2715 gctgttttcataaggttcctgccccctgacagacatcaaaaaatcagctcaagtcag 2774
 QY 576 AGGTGGCGAAACCGACAGGACTATAAAGATACCGAGGCTTTCCTGCTGAGCTCCCTC 635
 DB 2775 aggtggcgaaacccgacaggaactataaagataccagggcttccctcggagactccctc 2834
 QY 636 GTGGGCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGCTGCTGCTGCTGCTGCTG 695
 DB 2835 gtgctctctctgttcgacctgcccgttaccggatacctgtccgcttctccctctcg 2894
 QY 696 GGAAGCGTGGCGTTTCTCATAGCTACGCTGTAGGTATCTCA 738
 DB 2895 ggaagcgctggtttctcatagctacgctgtaggtatctca 2937

RESULT 13

AAV37292

XX AAV37292 standard; DNA; 4539 BP.

XX AAV37292;

DT 10-SEP-1998 (first entry)

XX pINVI plasmid used in the course of the invention.

XX Plasmid pINVI; reverse-splicing intron; group II intron;

XX exon binding site; domain V motif; branch site acceptor;

XX nucleophilic group; transesterification; phosphodiester bond;

XX autocatalytic Y-branched intron; reverse splicing reaction; ds.
 OS Synthetic.

XX Key Location/Qualifiers
 FT exon 969..1259
 FT /*tag= a
 FT /*note= "E3 exon"
 FT 1290..1559
 FT /*tag= b
 FT /*note= "E5 exon"

XX US5780272-A.

XX 14-JUL-1998.

XX 07-JUN-1995; 95US-0488015.

XX 10-SEP-1993; 93US-0119512.

XX (HARD) HARVARD COLLEGE.

XX Jarrell KA;

XX WPI; 1998-413060/35.

Reverse splicing construct containing fragments of autocatalytic
 introns - able to cleave and ligate discontinuous nucleic acid for
 generating new genes and e.g. ribozymes, libraries of enzymes and
 antibodies

Example 1; Columns 45-50; 56pp; English.

The present sequence represents plasmid pINVI which is used in the
 course of the invention. The specification describes a purified
 reverse-splicing intron which comprises a segment comprising a
 5'-part of a group II intron, including an exon binding site not
 naturally present in the intron and a second segment comprising a
 3'-part of a group II intron, including a domain V motif, a branch
 site acceptor, and a nucleophilic group for transesterifying a
 phosphodiester bond of an RNA. Together the two segments form an
 autocatalytic Y-branched intron which catalyses integration of at least
 the first segment into substrate RNA by a reverse splicing reaction
 The reverse-splicing introns are used, by specific cleavage and ligation
 of discontinuous nucleic acid, to generate new genes and gene products,
 e.g. ribozymes (for use in gene therapy or as reagents in DNA
 manipulation, e.g. replacements for restriction enzymes) or
 immunologically active or signal-transducing proteins such as antibody
 and enzyme libraries.

Sequence 4539 BP; 1246 A; 979 C; 998 G; 1316 T; 0 other;

Query Match 90.5%; Score 667.8; DB 19; Length 4539;
 Best Local Similarity 97.9%; Pred. No. 5.1e-112;
 Matches 688; Conservative 0; Mismatches 12; Indels 3; Gaps 1;

QY 39 GGGGATCCTCTAGAGTCGACCTGCGAGCATGCAAGCTTGAGTATTCTATAGTGTACACCA 98
 DB 2235 ggggattcctctagagtcgacctgcagcatgcaagctttgtcccttttagtgagggtta 2294
 QY 99 AAT---AGCTTGGCTTAATCATGTCATAGCTGTTTCTGTTGTAATTCCTGCTC 155
 DB 2295 atttcgagcttggtgtaatacattggtcattagctgtttcctgtggaattgttaccgctc 2354
 QY 156 ACAATTCCACACATAGCAGCGGAGCATATAAGTGTAAAGCCTGGGGTCCCTAATGA 215
 DB 2355 acaattccacacacatacagccggaagcataaagtgtaagcctgggggacctcaatga 2414
 QY 216 GTGAGTAACTCACATTAATTCGTTGCTCCTCACTGCCGCTTCCAGTCGGGAAACCTG 275
 DB 2415 gtgagctaaactacattaatgcttggtgctcactgcccgttccagctcgggaaacctg 2474

QY 276 TCGTCCAGCTGCATTAAATGATCGCCCAACCGCGGGAGAGCGGTTTCGTTATGGG 335
Db 2475 tctgtccagctgcatttaataatgaatcgcccaacgcgcggagagcggttgcgtattggg 2534
QY 336 CGCTCTCCGCTTCCTCGCTCAGTCAGTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 395
Db 2535 cgccttcctcgttcctcgtcactgaactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 2594
QY 396 GTATCAGCTCACTCAAGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACACGGA 455
Db 2595 gtatcagctcaactcaagcggtaatacgtttatccacagatacagggataacacagga 2654
QY 456 AAGACATGTCAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAATAAGCGCGCTGCTG 515
Db 2655 aagaacatgtgagcaaaagccagcaaaagccaggaacccgtataaagcgcgttgctg 2714
QY 516 GCGTTTTTCCATAGCTCCGCCCGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAG 575
Db 2715 gcgttttccatagctccgcccgccctgacgagcatcaaaaaatcgacgtcaagtca 2774
QY 576 AGGTGGCGAAACCCGACAGGACTATAAGATACCAAGGCGTTTCCCGCTGGAAGCTCCCTC 635
Db 2775 aggtggcgaaacccgacagactataaagataccagcggttccctccctggaagtcctc 2834
QY 636 GTGGCTCTCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCG 695
Db 2835 gtggctctcctgttccgacccctgctaccggatccgtgctcgtcgtcgtcgtcgtcgt 2894
QY 696 GGAAGCTGGCGCTTTCATAGCTACGCTGATAGGTATCTCA 738
Db 2895 ggaagcgtgcttctcatagctacgctgtaggtatctca 2937

RESULT 14
AAC62824/c
ID AAC62824 standard; DNA; 2571 BP.
XX AAC62824;
AC AAC62824;
XX AAC62824;
DT 02-FEB-2001 (first entry)
XX
DE Glyceraldehyde-3-phosphate dehydrogenase promoter coding sequence #3.
XX
XX Promoter; eucalyptus; pine; gene transcription; ds.
KW Promoter; eucalyptus; pine; gene transcription; ds.
OS Eucalyptus grandis.
XX
XX WO200058474-A1.
XX
XX 05-OCT-2000.
XX
XX 24-FEB-2000; 2000WO-NZ00018.
XX
XX 25-MAR-1999; 99US-0276599.
PR 30-JUL-1999; 99US-0146591.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
XX Perera R, Rice S, Eagleton CK;
PI WPI; 2000-647236/62.
DR
XX Novel promoter sequences useful for modulating transcription of plant
PT DNA sequences of interest and production of polypeptides -
XX
XX Claim 1; Pages 79-80; 93pp; English.
XX
CC The present invention relates to promoter sequences from eucalyptus
CC and pine. The present sequence is one such promoter. This sequence is
CC useful for modulating the transcription of DNA sequences of interest. The
CC sequences may also be used to tag or identify an organism or its
CC reproductive material.

XX
SQ Sequence 2571 BP; 659 A; 608 C; 635 G; 669 T; 0 other;
Query Match 90.4%; Score 667; DB 21; Length 2571;
Best Local Similarity 100.0%; Pred. No. 7.8e-112;
Matches 667; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 72 AGCTTGAGTATCTATAGTGTCACTAAATAGCTTGGCGTAATCATATGGTCATAGCTGTTT 131
Db 932 AGCTTGAGTATCTATAGTGTCACTAAATAGCTTGGCGTAATCATATGGTCATAGCTGTTT 873
QY 132 CCGTGTGAAATGTTATCCGCTCACAATTCACACAAATAGAGCCGGAAGCATAAAG 191
Db 872 CCGTGTGAAATGTTATCCGCTCACAATTCACACAAATAGAGCCGGAAGCATAAAG 813
QY 192 TGTAAAGCTGGGTCCTTAATAGTGTGAGTCACTCACTCACTCACTCACTCACTCACTCACTG 251
Db 812 TGTAAAGCTGGGTCCTTAATAGTGTGAGTCACTCACTCACTCACTCACTCACTCACTG 753
QY 252 CCCGCTTTCCAGTCCGGGAACCTGTCGTCGAGTCGATTAATGAATCGGCCAACCGCGG 311
Db 752 CCCGCTTTCCAGTCCGGGAACCTGTCGTCGAGTCGATTAATGAATCGGCCAACCGCGG 693
QY 312 GGGAGAGCGGTTTGGCTATTTGGCGCTCTTCCGCTTCTCGCTCACTGCTGCTGCGG 371
Db 692 GGGAGAGCGGTTTGGCTATTTGGCGCTCTTCCGCTTCTCGCTCACTGCTGCTGCGG 633
QY 372 TCGGTCGTTCCGCTCGGCGAGCGGTATCAGCTCACTCAAAAGCGGTAATACGTTTATCC 431
Db 632 TCGGTCGTTCCGCTCGGCGAGCGGTATCAGCTCACTCAAAAGCGGTAATACGTTTATCC 573
QY 432 ACAGAAATCAGGGGTAACGCAAGAAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGG 491
Db 572 ACAGAAATCAGGGGTAACGCAAGAAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGG 513
QY 492 AACCGTAAAGAGCGCGCTTGTGCGCTTTTCCATAGGCTCCGCCCTCCCTCAGCAGCAT 551
Db 512 AACCGTAAAGAGCGCGCTTGTGCGCTTTTCCATAGGCTCCGCCCTCCCTCAGCAGCAT 453
QY 552 CACAAAATCAGCGCTCAAGTCAGAGGTGGGAAACCGCACAGGACTATAAAGATACCAAG 611
Db 452 CACAAAATCAGCGCTCAAGTCAGAGGTGGGAAACCGCACAGGACTATAAAGATACCAAG 393
QY 612 GCGTTTCCCGCTTGAAGCTCCCTCGTGCCTCTCTGTTCCGACCTCGCGCTTACCGGA 671
Db 392 GCGTTTCCCGCTTGAAGCTCCCTCGTGCCTCTCTGTTCCGACCTCGCGCTTACCGGA 333
QY 672 TACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGG 731
Db 332 TACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGG 273
QY 732 TATCTCA 738
Db 272 TATCTCA 266

RESULT 15
ABK17089/c
ID ABK17089 standard; cDNA; 2571 BP.
XX ABK17089;
AC ABK17089;
XX
XX 26-MAR-2002 (first entry)
DT
XX Eucalyptus grandis promoter polynucleotide #43.
DE
XX Promoter; pine; leaf; root; flower; pollen; bud; meristem; xylogenesis;
KW temporally regulated promoter; Pinus radiata; Eucalyptus grandis; ss;
KW PCR primer.
XX
XX Eucalyptus grandis.
OS

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2002, 09:30:48 ; Search time 87.24 Seconds
(without alignments)
2077.918 Million cell updates/sec

Title: US-10-014-743-3
Perfect score: 738
Sequence: 1 ATACGACTACTAGGGCG.....CTCAGCGCTAGGTATCTCA 738

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues
Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTBUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	738	100.0	738	1	US-08-726-462B-3
2	738	100.0	738	2	US-09-046-203-3
3	738	100.0	738	3	US-09-272-104-3
4	738	100.0	738	4	US-09-272-097-3
5	672.4	91.1	4283	1	US-08-343-401A-3
6	672.4	91.1	4283	1	US-08-445-265A-1
7	672.4	91.1	4283	1	US-08-990-442-1
8	667.8	90.5	4539	1	US-08-119-512-1
9	667.8	90.5	4539	1	US-08-488-015B-1
10	667.8	90.5	4542	3	US-08-814-412-11
11	662.2	89.7	3485	4	US-08-939-366-10
12	660.4	89.5	10306	3	US-08-716-351A-4
13	660.4	89.5	10970	3	US-08-716-351A-5
14	659.4	89.3	9837	1	US-08-832-883-68
15	659.4	89.3	9837	2	US-08-832-877-68
16	650.8	88.2	4161	4	US-09-185-244-8
17	646.2	87.6	3318	2	US-08-495-500-2
18	641.6	86.9	8854	3	US-09-053-549-1
19	638.4	86.5	4824	2	US-08-485-139-5
20	638.4	86.5	4824	3	US-08-750-357-5
21	637.6	86.4	8710	1	US-08-480-882B-3
22	637.6	86.4	8710	1	US-08-480-210-3
23	637.6	86.4	9019	1	US-08-480-882B-4
24	637.6	86.4	9019	1	US-08-480-210-4
25	637	86.3	3699	3	US-08-646-538-6
26	637	86.3	3699	4	US-09-503-222-6
27	637	86.3	3737	2	US-08-784-208-1

c 28 637 86.3 4435 2 US-08-792-824-1
29 637 86.3 4965 3 US-08-675-566-22
30 637 86.3 5109 3 US-08-675-566-19
31 637 86.3 5452 2 US-09-130-114-1
32 637 86.3 5618 3 US-08-799-569-1
33 637 86.3 5653 1 US-08-073-836-3
34 637 86.3 5653 1 US-08-235-277-1
35 637 86.3 5707 2 US-08-472-809B-8
36 637 86.3 6045 3 US-08-675-566-18
37 637 86.3 6196 3 US-08-675-566-5
38 637 86.3 6243 3 US-08-675-566-14
39 637 86.3 6243 3 US-08-675-566-17
40 637 86.3 6295 2 US-08-659-206A-4
41 637 86.3 6345 1 US-08-472-809B-7
42 637 86.3 6367 1 US-08-470-299-1
43 637 86.3 6367 3 US-08-776-511-3
44 637 86.3 6447 3 US-08-675-566-16
45 637 86.3 6503 3 US-08-675-566-6

ALIGNMENTS

RESULT 1
US-08-726-462B-3
; Sequence 3, Application US/08726462B
; Patent No. 5800996
; GENERAL INFORMATION:
; APPLICANT: Perkin-Elmer Corporation, Applied Biosystems
; APPLICANT: Division
; TITLE OF INVENTION: ENERGY TRANSFER DYES WITH ENHANCED
; TITLE OF INVENTION: FLUORESCENCE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David J. Weitz, Wilson Sonsini Goodrich
; ADDRESS: 650 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1050
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: Microsoft Windows 3.1/DOS 5.0
; SOFTWARE: Wordperfect for windows 6.0,
; SOFTWARE: ASCII (DOS) TEXT format
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08726,462B
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/642,330
; FILING DATE: May 3, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/672,196
; FILING DATE: June 27, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: David J. Weitz
; REGISTRATION NUMBER: 38,362
; REFERENCE/DOCKET NUMBER: PELM4304
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 493-9300
; TELEFAX: (415) 493-6811
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 738 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-726-462B-3

Sequence 1, Appl
Sequence 22, Appl
Sequence 19, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 3, Appl
Sequence 1, Appl
Sequence 8, Appl
Sequence 18, Appl
Sequence 5, Appl
Sequence 14, Appl
Sequence 17, Appl
Sequence 4, Appl
Sequence 7, Appl
Sequence 1, Appl
Sequence 3, Appl
Sequence 16, Appl
Sequence 6, Appl

QY 481 AAAAGGCCAGGACCGTAAAAAGGCCGGTTCCTGGCGTTTTCATAGGCTCCGCCCC 540
Db 481 AAAAGGCCAGGACCGTAAAAAGGCCGGTTCCTGGCGTTTTCATAGGCTCCGCCCC 540
QY 541 CTGACGAGCATCAAAAATCGAGCTCAAGTCAAGAGTGGCGAACCAGAGGACTAT 600
Db 541 CTGACGAGCATCAAAAATCGAGCTCAAGTCAAGAGTGGCGAACCAGAGGACTAT 600
QY 601 AAAGATACCGAGCGTTTCCCGCTGGAAGCTCCCTGCGTCTCTCTCTCTCTCTCT 660
Db 601 AAAGATACCGAGCGTTTCCCGCTGGAAGCTCCCTGCGTCTCTCTCTCTCTCTCT 660
QY 661 CGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGAAGCTGGCGTCTCTCTCTCT 720
Db 661 CGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGAAGCTGGCGTCTCTCTCTCT 720
QY 721 CACGCTGTAGTATCTCA 738
Db 721 CACGCTGTAGTATCTCA 738

RESULT 3
US-09-272-104-3
; Sequence 3, Application US/09272104
; Patent No. 6145434
; GENERAL INFORMATION:
; APPLICANT: Perkin-Elmer Corporation, Applied Biosystems
; APPLICANT: Division
; TITLE OF INVENTION: METHOD FOR DETECTING
; TITLE OF INVENTION: OLIGONUCLEOTIDES USING ENERGY TRANSFER DYES WITH LONG STROKE SH
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David J. Weitz, Wilson Sonsini Goodrich
; STREET: 650 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1050
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: Microsoft Windows 3.1/DOS 5.0
; SOFTWARE: Wordperfect for windows 6.0,
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/272,104
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/642,330
; FILING DATE: May 3, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/672,196
; FILING DATE: June 27, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/726,452
; FILING DATE: October 4, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/046,203
; FILING DATE: March 23, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: David J. Weitz
; REGISTRATION NUMBER: 38,362
; REFERENCE/DOCKET NUMBER: 16842-755
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 493-9300
; TELEFAX: (415) 493-6811
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 738 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; US-09-272-104-3
Query Match 100.0%; Score 738; DB 3; Length 738;
Best Local Similarity 100.0%; Pred. No. 1.6e-175;
Matches 738; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATACGACTCAGTATAGGGGAAATTCGAGTTCGAGTCCCGGGGATCTCTCTAGAGTCGACCT 60
Db 1 ATACGACTCAGTATAGGGGAAATTCGAGTTCGAGTCCCGGGGATCTCTCTAGAGTCGACCT 60
QY 61 GCAGGCATCGAAGCTTGAGTATTCATAGTGTACCTTAATAGTCTGGCGTAATCATGCT 120
Db 61 GCAGGCATCGAAGCTTGAGTATTCATAGTGTACCTTAATAGTCTGGCGTAATCATGCT 120
QY 121 CATAGCTGTTTCTGTTGAAATTTGTTATCCGCTCACAATTTCCACACACATACGAGCCG 180
Db 121 CATAGCTGTTTCTGTTGAAATTTGTTATCCGCTCACAATTTCCACACACATACGAGCCG 180
QY 181 GAAGCATAAAGTGTAAAGCTTGGGTGCTTAATGAGTGAAGTAACTACATTAATTCGCT 240
Db 181 GAAGCATAAAGTGTAAAGCTTGGGTGCTTAATGAGTGAAGTAACTACATTAATTCGCT 240
QY 241 TCGCGTCACTGCCCGCTTTCAGTTCGGGAAACCTGTCGTCGCGGATTAATGATCG 300
Db 241 TCGCGTCACTGCCCGCTTTCAGTTCGGGAAACCTGTCGTCGCGGATTAATGATCG 300
QY 301 GCCAACGCGCGGGAGAGCGGTTTCCGTTATTTGGGCGCTTCCGCTTCCGCTCAGCTG 360
Db 301 GCCAACGCGCGGGAGAGCGGTTTCCGTTATTTGGGCGCTTCCGCTTCCGCTCAGCTG 360
QY 361 ACTCGCTCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 361 ACTCGCTCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 421 TACGCTTATCCACAGAAATCAGGGGATACGCGAGGAGGAAACATGTGAGCAAAAGGCCAGC 480
Db 421 TACGCTTATCCACAGAAATCAGGGGATACGCGAGGAGGAAACATGTGAGCAAAAGGCCAGC 480
QY 481 AAAAGGCCAGGAAACCGTAAAAAGGCCGGTTCCTGCGCTTTTCCATAGGCTCCGCCCC 540
Db 481 AAAAGGCCAGGAAACCGTAAAAAGGCCGGTTCCTGCGCTTTTCCATAGGCTCCGCCCC 540
QY 541 CTGACGAGCATCAAAAATCGAGCTCAAGTCAAGAGTGGCGAACCAGAGGACTAT 600
Db 541 CTGACGAGCATCAAAAATCGAGCTCAAGTCAAGAGTGGCGAACCAGAGGACTAT 600
QY 601 AAAGATACCGAGCGTTTCCCGCTGGAAGCTCCCTGCGTCTCTCTCTCTCTCTCTCT 660
Db 601 AAAGATACCGAGCGTTTCCCGCTGGAAGCTCCCTGCGTCTCTCTCTCTCTCTCTCT 660
QY 661 CGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGAAGCTGGCGTCTCTCTCTCTCT 720
Db 661 CGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGAAGCTGGCGTCTCTCTCTCTCT 720
QY 721 CACGCTGTAGTATCTCA 738
Db 721 CACGCTGTAGTATCTCA 738

RESULT 4
US-09-272-097-3
; Sequence 3, Application US/09272097
; Patent No. 633440
; GENERAL INFORMATION:
; APPLICANT: PE Corporation (NY)
; TITLE OF INVENTION: METHOD FOR DETECTING OLIGONUCLEOTIDES USING
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David J. Weitz, Wilson Sonsini Goodrich
; & Rosati

```

; STREET: 650 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1050
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; OPERATING SYSTEM: Microsoft Windows 3.1/DOS 5.0
; SOFTWARE: Word 97
; ASCII (DOS) TEXT format
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/272,097
; FILING DATE: 18-Mar-1999
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/642,330
; FILING DATE: May 3, 1996
; APPLICATION NUMBER: 08/672,196
; FILING DATE: June 27, 1996
; APPLICATION NUMBER: 08/726,462
; FILING DATE: October 4, 1996
; APPLICATION NUMBER: 09/046,203
; FILING DATE: March 23, 1998
;
; ATTORNEY/AGENT INFORMATION:
; NAME: David J. Weitz
; REGISTRATION NUMBER: 38,362
; REFERENCE/DOCKET NUMBER: 16842-755
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 493-9300
; TELEFAX: (650) 493-6811
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 738 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-272-097-3

```

```

Query Match      100.0%; Score 738; DB 4; Length 738;
Best Local Similarity 100.0%; Pred. No. 1.6e-175;
Matches 738; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAGGACTACTATAGGGCGAATTCGAGCTCGGTACCCGGGATCCCTCTAGAGTCGACCT 60
DB 1 ATAGGACTACTATAGGGCGAATTCGAGCTCGGTACCCGGGATCCCTCTAGAGTCGACCT 60

QY 61 GCAGGCGATGCAAGCTTGAGTATTCTATAGTGTACCTAAATAGCTTGGCGTAAATCATGGT 120
DB 61 GCAGGCGATGCAAGCTTGAGTATTCTATAGTGTACCTAAATAGCTTGGCGTAAATCATGGT 120

QY 121 CATAGCTGTTTCTGTGTAATTTGTTATCGCTCACAATTTCCACAAACATACGAGCG 180
DB 121 CATAGCTGTTTCTGTGTAATTTGTTATCGCTCACAATTTCCACAAACATACGAGCG 180

QY 181 GAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAATGAGTGAGTAACTACATTAATTCGGT 240
DB 181 GAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAATGAGTGAGTAACTACATTAATTCGGT 240

QY 241 TGGCTCACTGCGCGCTTTCAGTCGCGGAACCTGTGTCGCGCTTCCCTCGCTCACTG 300
DB 241 TGGCTCACTGCGCGCTTTCAGTCGCGGAACCTGTGTCGCGCTTCCCTCGCTCACTG 300

QY 301 GCCAAGCGGGGAGAGGGGTTTGGGTATTGGCGCTTTCGCGCTTCCCTCGCTCACTG 360
DB 301 GCCAAGCGGGGAGAGGGGTTTGGGTATTGGCGCTTTCGCGCTTCCCTCGCTCACTG 360

QY 361 ACTCGCTCGCGCTCGGTGCTTGGGTGCGGCGAGCGGTATCAGTCACTCAAAAGCGGTAA 420
DB 361 ACTCGCTCGCGCTCGGTGCTTGGGTGCGGCGAGCGGTATCAGTCACTCAAAAGCGGTAA 420

```

```

RESULT 5
US-08-343-401A-3
; Sequence 3, Application US/08343401A
; Patent No. 5661132
; GENERAL INFORMATION:
; APPLICANT: Swain, William F
; APPLICANT: Macklin, Michael D
; APPLICANT: Eriksson, Eloy
; APPLICANT: Andree, Christophe
; TITLE OF INVENTION: Improved Wound Healing
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Quares & Brady
; STREET: PO Box 2113
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,401A
; FILING DATE: 22-NOV-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 11-229-9103-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4283 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: pWRG1630
; FEATURE:
; NAME/KEY: exon
; LOCATION: 713..721
; NAME/KEY: exon

```

```

QY 421 TACGGTTATCCACAGAAATCAGGGGATTAACGAGGAAAGAACATCTGAGCAAAAGCCAGC 480
DB 421 TACGGTTATCCACAGAAATCAGGGGATTAACGAGGAAAGAACATCTGAGCAAAAGCCAGC 480

QY 481 AAAAGGCCAGGAACCGTAAAGGCGGTTGGCTGGCGTTTTCATAGGCTCCGCCCCC 540
DB 481 AAAAGGCCAGGAACCGTAAAGGCGGTTGGCTGGCGTTTTCATAGGCTCCGCCCCC 540

QY 541 CTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGACTAT 600
DB 541 CTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGACTAT 600

QY 601 AAGATACCGAGCGCTTCCCGCTGGAAGCTCCCTCGCGCTCTCCTGTCGACCCCTGC 660
DB 601 AAGATACCGAGCGCTTCCCGCTGGAAGCTCCCTCGCGCTCTCCTGTCGACCCCTGC 660

QY 661 CGCTTACCGGATACCTGTCCGCCCTTTCCTCCCTTCGGGAAGCGTGGCGCTTTCATAGCT 720
DB 661 CGCTTACCGGATACCTGTCCGCCCTTTCCTCCCTTCGGGAAGCGTGGCGCTTTCATAGCT 720

QY 721 CACGCTGTAGTATCTCA 738
DB 721 CACGCTGTAGTATCTCA 738

```


LOCATION: 981...1253
FEATURE: CDS
NAME/KEY: join(713...721, 981...1253)
FEATURE: sig_peptide
LOCATION: 713...1049
US-08-343-401A-3

Query Match 91.1%; Score 672.4; DB 1; Length 4283;
Best Local Similarity 99.9%; Pred. No. 5e-159;
Matches 673; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 65 GCATGCAAGCTTGAGTATCTATAGTGCACCTAAATAGCTGGCGTAATCATGTCATA 124
DB 1590 GCATGCAAGCTTGAGTATCTATAGTGCACCTAAATAGCTGGCGTAATCATGTCATA 1649
QY 125 GCTGTTTCTGTTGTAATGTTTATCCGCTCACAATTCACACACATACGAGCGGAAG 184
DB 1650 GCTGTTTCTGTTGTAATGTTTATCCGCTCACAATTCACACACATACGAGCGGAAG 1709
QY 185 CATAAAGTGAAGCCCTGGGTCCTATAGTGCACCTAAATAGCTGGCGTAATCATGTCATA 244
DB 1710 CATAAAGTGAAGCCCTGGGTCCTATAGTGCACCTAAATAGCTGGCGTAATCATGTCATA 1769
QY 245 CTCACTGCGGCTTCCAGTCGGGAAACCTGCTGCGGAGCGGTATCACTCAAAAGCGGTAATACG 304
DB 1770 CTCACTGCGGCTTCCAGTCGGGAAACCTGCTGCGGAGCGGTATCACTCAAAAGCGGTAATACG 1829
QY 305 ACGCGGGGAGAGCGGTTGCGTATTTGGCGCTCTTCCGCTTCCGCTCACTGACTC 364
DB 1830 ACGCGGGGAGAGCGGTTGCGTATTTGGCGCTCTTCCGCTTCCGCTCACTGACTC 1889
QY 365 GCTGCGCTCGGTCGTTCCGCTGCGGAGCGGTATCACTCAAAAGCGGTAATACG 424
DB 1890 GCTGCGCTCGGTCGTTCCGCTGCGGAGCGGTATCACTCAAAAGCGGTAATACG 1949
QY 425 GTTATCCACAGAAATCAGGAGTAACGAGGAAACATGAGCAAAAGCGGTAATACG 484
DB 1950 GTTATCCACAGAAATCAGGAGTAACGAGGAAACATGAGCAAAAGCGGTAATACG 2009
QY 485 GCGCAGAAACCGTAAAGGCGGTTGCTGGCGTCTTCCATAGGTCGCGCCCGCTGA 544
DB 2010 GCGCAGAAACCGTAAAGGCGGTTGCTGGCGTCTTCCATAGGTCGCGCCCGCTGA 2069
QY 545 CGAGCATCAAAAATCGAGCTCAAGTCAGAGTGCGGAAACCGCAGAGCTATAAG 604
DB 2070 CGAGCATCAAAAATCGAGCTCAAGTCAGAGTGCGGAAACCGCAGAGCTATAAG 2129
QY 605 ATACGAGCGGTTCCCGCTGGAAGCTCCCTGCTGCGCTCTCTGCTGCGAGCCCTGCGCT 664
DB 2130 ATACGAGCGGTTCCCGCTGGAAGCTCCCTGCTGCGCTCTCTGCTGCGAGCCCTGCGCT 2189
QY 665 TACGGATACCTGTCGCGCTTCTCCCTTCGGGAAAGCGTGGCGCTTCTCATAGCTACG 724
DB 2190 TACGGATACCTGTCGCGCTTCTCCCTTCGGGAAAGCGTGGCGCTTCTCATAGCTACG 2249
QY 725 CTGTAGTATCTCA 738
DB 2250 CTGTAGTATCTCA 2263

RESULT 6
US-08-445-265A-1
Sequence 1, Application US/08445265A
Patent No. 5697901
GENERAL INFORMATION:
APPLICANT: Eriksson, Elof
TITLE OF INVENTION: GENE DELIVERY BY MICRONEEDLE INJECTION
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady

STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI US
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/445,265A
FILING DATE:
CLASSIFICATION: 604
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 110229.91080
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4283 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Plasmid DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: join(713...721, 981...1250)
US-08-445-265A-1

Query Match 91.1%; Score 672.4; DB 1; Length 4283;
Best Local Similarity 99.9%; Pred. No. 5e-159;
Matches 673; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 65 GCATGCAAGCTTGAGTATCTATAGTGCACCTAAATAGCTGGCGTAATCATGTCATA 124
DB 1590 GCATGCAAGCTTGAGTATCTATAGTGCACCTAAATAGCTGGCGTAATCATGTCATA 1649
QY 125 GCTGTTTCTGTTGTAATGTTTATCCGCTCACAATTCACACACATACGAGCGGAAG 184
DB 1650 GCTGTTTCTGTTGTAATGTTTATCCGCTCACAATTCACACACATACGAGCGGAAG 1709
QY 185 CATAAAGTGAAGCCCTGGGTCCTATAGTGCACCTAAATAGCTGGCGTAATCATGTCATA 244
DB 1710 CATAAAGTGAAGCCCTGGGTCCTATAGTGCACCTAAATAGCTGGCGTAATCATGTCATA 1769
QY 245 CTCACTGCGGCTTCCAGTCGGGAAACCTGCTGCGGAGCGGTATCACTCAAAAGCGGTAATACG 304
DB 1770 CTCACTGCGGCTTCCAGTCGGGAAACCTGCTGCGGAGCGGTATCACTCAAAAGCGGTAATACG 1829
QY 305 ACGCGGGGAGAGCGGTTGCGTATTTGGCGCTCTTCCGCTTCCGCTCACTGACTC 364
DB 1830 ACGCGGGGAGAGCGGTTGCGTATTTGGCGCTCTTCCGCTTCCGCTCACTGACTC 1889
QY 365 GCTGCGCTCGGTCGTTCCGCTGCGGAGCGGTATCACTCAAAAGCGGTAATACG 424
DB 1890 GCTGCGCTCGGTCGTTCCGCTGCGGAGCGGTATCACTCAAAAGCGGTAATACG 1949
QY 425 GTTATCCACAGAAATCAGGAGTAACGAGGAAACATGAGCAAAAGCGGTAATACG 484
DB 1950 GTTATCCACAGAAATCAGGAGTAACGAGGAAACATGAGCAAAAGCGGTAATACG 2009
QY 485 GCGCAGAAACCGTAAAGGCGGTTGCTGGCGTCTTCCATAGGTCGCGCCCGCTGA 544
DB 2010 GCGCAGAAACCGTAAAGGCGGTTGCTGGCGTCTTCCATAGGTCGCGCCCGCTGA 2069
QY 545 CGAGCATCAAAAATCGAGCTCAAGTCAGAGTGCGGAAACCGCAGAGCTATAAG 604
DB 2070 CGAGCATCAAAAATCGAGCTCAAGTCAGAGTGCGGAAACCGCAGAGCTATAAG 2129
QY 605 ATACGAGCGGTTCCCGCTGGAAGCTCCCTGCTGCGCTCTCTGCTGCGAGCCCTGCGCT 664
DB 2130 ATACGAGCGGTTCCCGCTGGAAGCTCCCTGCTGCGCTCTCTGCTGCGAGCCCTGCGCT 2189
QY 665 TACGGATACCTGTCGCGCTTCTCCCTTCGGGAAAGCGTGGCGCTTCTCATAGCTACG 724
DB 2190 TACGGATACCTGTCGCGCTTCTCCCTTCGGGAAAGCGTGGCGCTTCTCATAGCTACG 2249
QY 725 CTGTAGTATCTCA 738
DB 2250 CTGTAGTATCTCA 2263


```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 4539 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: other nucleic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 969..1259
; OTHER INFORMATION: /product= "E3 exon"
;
; NAME/KEY: misc_feature
; LOCATION: 1290..1559
; OTHER INFORMATION: /product= "E5 exon"
;
US-08-119-512-1

```

```

Query Match          90.5%; Score 667.8; DB 1; Length 4539;
Best Local Similarity 97.9%; Pred. No. 7e-158;
Matches 688; Conservative 0; Mismatches 12; Indels 3; Gaps 1;

QY 39 GGGATCCTCTAGAGTCGACCTGAGCGATGCAAGCTTGTAGTATTTCTATAGTGTCACTA 98
DB 2235 GGGATCCTCTAGAGTCGACCTGAGCGATGCAAGCTTGTAGTATTTCTATAGTGTCA 2234
QY 99 AAT---AGCTGGCGTAATCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 155
DB 2295 ATTTTCGAGCTTGGCGTAATCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2354
QY 156 ACAATTCACACACATAGAGCGGAGCAATAAGTCTAAAGCTGGGTCGCTTAATGA 215
DB 2355 ACAATTCACACACATAGAGCGGAGCAATAAGTCTAAAGCTGGGTCGCTTAATGA 2414
QY 216 GTGAGTAACTCACTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 275
DB 2415 GTGAGTAACTCACTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2474
QY 276 TCGTGGCGAGTCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 335
DB 2475 TCGTGGCGAGTCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2534
QY 336 CGCTCTTCGCTTCTCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 395
DB 2535 CGCTCTTCGCTTCTCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2594
QY 396 GTATCAGTCACTCAAGCGGTAATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 455
DB 2595 GTATCAGTCACTCAAGCGGTAATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2654
QY 456 AAGAACATGTGAGCAAAAGCGGAGCAAAAGCGGAGCAAAAGCGGAGCAAAAGCGGAG 515
DB 2655 AAGAACATGTGAGCAAAAGCGGAGCAAAAGCGGAGCAAAAGCGGAGCAAAAGCGGAG 2714
QY 516 GCGTTTTCATAGGCTCCGCCCCCTGACGAGCATCAAAAGCTGAGGCTCAAGTCA 575
DB 2715 GCGTTTTCATAGGCTCCGCCCCCTGACGAGCATCAAAAGCTGAGGCTCAAGTCA 2774
QY 576 AGTGGCGAAGCCGACAGGACTATAAGTACCGAGCTTCCCGCTGGAAGCTCCCTC 635
DB 2775 AGTGGCGAAGCCGACAGGACTATAAGTACCGAGCTTCCCGCTGGAAGCTCCCTC 2834
QY 636 GTGCGCTCTCTGCTCCGAGCTTACCGGCTTACCGGATGCTGCTGCTGCTGCTGCTG 695
DB 2835 GTGCGCTCTCTGCTCCGAGCTTACCGGCTTACCGGATGCTGCTGCTGCTGCTGCTG 2894
QY 696 GGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGTATCTCA 738
DB 2895 GGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGTATCTCA 2937

```

```

RESULT 9
US-08-488-015B-1
; Sequence 1, Application US/08488015B

```

```

; Patent No. 5780272
; GENERAL INFORMATION:
; APPLICANT: Jarrell, Kevin A.
; TITLE OF INVENTION: INTRON-MEDIATED RECOMBINANT TECHNIQUES
; TITLE OF INVENTION: AND REAGENTS
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Elliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,015B
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HUV-008.02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4539 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: other nucleic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 969..1259
; OTHER INFORMATION: /product= "E3 exon"
;
; NAME/KEY: misc_feature
; LOCATION: 1290..1559
; OTHER INFORMATION: /product= "E5 exon"
;
US-08-488-015B-1

```

```

Query Match          90.5%; Score 667.8; DB 1; Length 4539;
Best Local Similarity 97.9%; Pred. No. 7e-158;
Matches 688; Conservative 0; Mismatches 12; Indels 3; Gaps 1;

QY 39 GGGATCCTCTAGAGTCGACCTGAGCGATGCAAGCTTGTAGTATTTCTATAGTGTCACTA 98
DB 2235 GGGATCCTCTAGAGTCGACCTGAGCGATGCAAGCTTGTAGTATTTCTATAGTGTCA 2294
QY 99 AAT---AGCTGGCGTAATCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 155
DB 2295 ATTTTCGAGCTTGGCGTAATCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2354
QY 156 ACAATTCACACACATAGAGCGGAGCAATAAGTCTAAAGCTGGGTCGCTTAATGA 215
DB 2355 ACAATTCACACACATAGAGCGGAGCAATAAGTCTAAAGCTGGGTCGCTTAATGA 2414
QY 216 GTGAGTAACTCACTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 275
DB 2415 GTGAGTAACTCACTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2474
QY 276 TCGTGGCGAGTCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 335
DB 2475 TCGTGGCGAGTCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2534
QY 336 CGCTCTTCGCTTCTCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 395
DB 2535 CGCTCTTCGCTTCTCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2594

```

QY 396 GTATCAGCTCACTCAAGGCGGTAAATACGGTTATCCACAGAAATCAAGGGATCAACGACGA 455
DB 2595 GTATCAGCTCACTCAAGGCGGTAAATACGGTTATCCACAGAAATCAAGGGATCAACGACGA 2654
QY 456 AAGAACATGTGAGCAAAAGGCGCAAGGCGGCAAGGCGGCAAGGCGGCGGTTGCTG 515
DB 2655 AAGAACATGTGAGCAAAAGGCGCAAGGCGGCAAGGCGGCGGTTGCTG 2714
QY 516 GCGTTTTCATAGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAG 575
DB 2715 GCGTTTTCATAGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAG 2774
QY 576 AGTGGGGAACCGACAGGACTATAAAGATACAGGCGTTTCCGCTTGGAGCTCCCTC 635
DB 2775 AGTGGGGAACCGACAGGACTATAAAGATACAGGCGTTTCCGCTTGGAGCTCCCTC 2834
QY 636 GTGCGCTCTCCTGTTCGACCCCTGACGAGTACCGGTTTCCGCTTGGAGCTCCCTC 2894
DB 2835 GTGCGCTCTCCTGTTCGACCCCTGACGAGTACCGGTTTCCGCTTGGAGCTCCCTC 2937
QY 696 GGAAGCGTGGCGCTTCTCATAGCTACGCTGAGGTATCTCA 738
DB 2895 GGAAGCGTGGCGCTTCTCATAGCTACGCTGAGGTATCTCA 2937

RESULT 10
US-08-814-412-11
; Sequence 11, Application US/08814412
; Patent No. 6150141
; GENERAL INFORMATION:
; APPLICANT: Jarrell Ph.D., Kevin A.
; TITLE OF INVENTION: Intron-Mediated Techniques and Reagents
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Choate, Hall & Stewart
; STREET: 53 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,412
; FILING DATE: 11-MAR-1997
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
; NAME: Jarrell Ph.D., Brenda H.
; REGISTRATION NUMBER: 39,223
; REFERENCE/DOCKET NUMBER: 0079571-0040
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 248 5000
; TELEFAX: 617 248 4000
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4542 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "plasmid DNA"
; IMMEDIATE SOURCE:
; CLONE: pINVL
; US-08-814-412-11

Query Match 90.5%; Score 667.8; DB 3; Length 4542;
Best Local Similarity 97.9%; Pred. No. 7e-158;
Matches 688; Conservative 0; Mismatches 12; Indels 3; Gaps 1;

QY 39 GGGATCCCTCTAGAGTCGACCTCGACGATCGAAGCTTGTAGTATCTTATAGTGTACCTA 98
DB 2238 GGGATCCCTCTAGAGTCGACCTCGACGATCGAAGCTTGTAGTATCTTATAGTGTACCTA 2297
QY 99 AAT---AGCTTGGCGTAATCATGGTCATAGCTGTTTCTCTGTGAAATCTTATCCGCTC 155
DB 2298 AATTCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCTCTGTGAAATCTTATCCGCTC 2357
QY 156 ACAATTCACACAACATACAGCGCGGAAGCATAAAGTGTAAAGCTTGGGTCGCTAATGA 215
DB 2358 ACAATTCACACAACATACAGCGCGGAAGCATAAAGTGTAAAGCTTGGGTCGCTAATGA 2417
QY 216 GTGAGCTAACTCACATTAATTCGCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTG 275
DB 2418 GTGAGCTAACTCACATTAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2477
QY 276 TGGTCCAGCTCACTCAAAAGCGGTAAATACGCTTATCCACAGAAATCAAGGGATCAACGACGA 455
DB 2478 TGGTCCAGCTCACTCAAAAGCGGTAAATACGCTTATCCACAGAAATCAAGGGATCAACGACGA 2657
QY 336 CGCTCTTCCGCTTCCCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 395
DB 2538 CGCTCTTCCGCTTCCCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2597
QY 396 GTATCAGCTCACTCAAAAGCGGTAAATACGCTTATCCACAGAAATCAAGGGATCAACGACGA 455
DB 2598 GTATCAGCTCACTCAAAAGCGGTAAATACGCTTATCCACAGAAATCAAGGGATCAACGACGA 2657
QY 456 AAGAACATGTGAGCAAAAGGCGCAAGGCGGCAAGGCGGCAAGGCGGCGGTTGCTG 515
DB 2658 AAGAACATGTGAGCAAAAGGCGCAAGGCGGCAAGGCGGCAAGGCGGCGGTTGCTG 2717
QY 516 GCGTTTTCATAGCTCCGCCCCCTGACGAGTACCGGTTTCCGCTTGGAGCTCCCTC 575
DB 2718 GCGTTTTCATAGCTCCGCCCCCTGACGAGTACCGGTTTCCGCTTGGAGCTCCCTC 2777
QY 576 AGTGGGGAACCGACAGGACTATAAAGATACAGGCGTTTCCGCTTGGAGCTCCCTC 635
DB 2778 AGTGGGGAACCGACAGGACTATAAAGATACAGGCGTTTCCGCTTGGAGCTCCCTC 2837
QY 636 GTGCGCTCTCCTGTTCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCTCCCTTCG 695
DB 2838 GTGCGCTCTCCTGTTCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCTCCCTTCG 2897
QY 696 GGAAGCGTGGCGCTTCTCATAGCTACGCTGAGGTATCTCA 738
DB 2898 GGAAGCGTGGCGCTTCTCATAGCTACGCTGAGGTATCTCA 2940

RESULT 11
US-08-939-366-10
; Sequence 10, Application US/08939366
; Patent No. 6355415
; GENERAL INFORMATION:
; APPLICANT: Wagner, Thomas E.
; TITLE OF INVENTION: Compositions and Methods for the Use of
; TITLE OF INVENTION: Ribozymes to Determine Gene Function
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

Db 9946 ATCAGGGGATACGACAGAAAGAACATGTGAGCAAAAGGCCACAAAGGCCAGGACCG 9887
QY 497 TAAAGAGCGCGCTTCTGCGCGTTTTCATAGGCTCCGCCGCCCTTCACGAGCATCACAA 556
Db 9886 TAAAGAGCGCGCTTCTGCGCGTTTTCATAGGCTCCGCCGCCCTTCACGAGCATCACAA 9827
QY 557 AATCAGCGCTCAAGTCAGAGGTGGCGAAACCCGACGAGACTATAAGATACACAGCGTT 616
Db 9826 AATCAGCGCTCAAGTCAGAGGTGGCGAAACCCGACGAGACTATAAGATACACAGCGTT 9767
QY 617 TCCCGCTGGAAGCTCCCTCGTGGCGTCTCCTGTTCCGACCTCGCGCTTACCGATACCT 676
Db 9766 TCCCGCTGGAAGCTCCCTCGTGGCGTCTCCTGTTCCGACCTCGCGCTTACCGATACCT 9707
QY 677 GTCCGCTTCTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTCAGCGTGTAGGTATCT 736
Db 9706 GTCCGCTTCTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTCAGCGTGTAGGTATCT 9647
QY 737 CA 738
Db 9646 CA 9645

RESULT 13

US-08-716-351A-5/c
; Sequence 5, Application US/08716351A
; Patent No. 6033905
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Gibbon Ape Leukemia Virus-Based
; TITLE OF INVENTION: Retroviral Vectors
; NUMBER OF SEQUENCES: 5
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/716,351A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/03784
; FILING DATE: 06-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 15280-128-1PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10970 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..10970
; OTHER INFORMATION: /standard_name= "p537 retroviral"
; OTHER INFORMATION: vector"
US-08-716-351A-5

Query Match 89.5%; Score 660.4; DB 3; Length 10970;
Best Local Similarity 99.8%; Pred. No. 5.7e-156;
Matches 661; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 77 GAGTATTCATAGCTCACCTAAATAGCTGGCGTAATCATGCTCATAGCTGTTCTCTGT 136
|||||

Db 10970 GAGTATTCATAGCTCACCTAAATAGCTTGGCGTAATCATGCTCATAGCTGTTCTCTGT 10911
QY 137 GTGAATTTGTTATCCGCTCAACAATTCACACAAACATACGAGCCGGAAGCATAAAGTGAA 196
Db 10910 GTGAATTTGTTATCCGCTCAACAATTCACACAAACATACGAGCCGGAAGCATAAAGTGAA 10851
QY 197 AGCCTGGGTGCTTAATGAGTGAGCTAACTCACATTAATTCGGTTTGGCTCACTGCCCGC 256
Db 10850 AGCCTGGGTGCTTAATGAGTGAGCTAACTCACATTAATTCGGTTTGGCTCACTGCCCGC 10791
QY 257 TTTCCAGTTCGGGAAACCTGCTGCCAGCTCATTTAAATGAATCGGCCAAACGCGGGGAG 316
Db 10790 TTTCCAGTTCGGGAAACCTGCTGCCAGCTCATTTAAATGAATCGGCCAAACGCGGGGAG 10731
QY 317 AGCGGTTTGGGTATTTGGCGCTCTTCCGCTTCCGCTCACTCACTCGCTCGCTCGGT 376
Db 10730 AGCGGTTTGGGTATTTGGCGCTCTTCCGCTTCCGCTCACTCACTCGCTCGCTCGGT 10671
QY 377 CGTTCCGCTTCGGCGAGCGGTATCAGCTCACTCAAGGCGGTAAATACGGTTATCCACAGA 436
Db 10670 CGTTCCGCTTCGGCGAGCGGTATCAGCTCACTCAAGGCGGTAAATACGGTTATCCACAGA 10611
QY 437 ATCAGGGGATAACGACGAGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGAACCG 496
Db 10610 ATCAGGGGATAACGACGAGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGAACCG 10551
QY 497 TAAAGAGCGCGCTTCTGCTGGCGTTTTTCCATAGGCTCCGCCCTTCCGCTTACCGGATACCT 516
Db 10550 TAAAGAGCGCGCTTCTGCTGGCGTTTTTCCGATAGGCTCCGCCCTTCCGCTTACCGGATACCT 10431
QY 557 AATCAGGCTCAAGTCAGAGGTGGCGAAACCCGACAGACTATAAGATACCGAGCT 616
Db 10490 AATCAGGCTCAAGTCAGAGGTGGCGAAACCCGACAGACTATAAGATACCGAGCT 10431
QY 617 TCCCGCTGGAAGCTCCCTCGTGGCGTCTCCTGTTCCGACCTCGCGCTTACCGGATACCT 676
Db 10430 TCCCGCTGGAAGCTCCCTCGTGGCGTCTCCTGTTCCGACCTCGCGCTTACCGGATACCT 10371
QY 677 GTCCGCTTCTCCCTTCGGGAAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCT 736
Db 10370 GTCCGCTTCTCCCTTCGGGAAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCT 10311
QY 737 CA 738
Db 10310 CA 10309

RESULT 14

US-08-832-883-68/c
; Sequence 68, Application US/08832883
; Patent No. 5807681
; GENERAL INFORMATION:
; APPLICANT: Giordano, Antonio
; APPLICANT: Baldi, Alphonso
; TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND PROGNOSIS
; TITLE OF INVENTION: OF CANCER
; NUMBER OF SEQUENCES: 115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
; STREET: Suite 1800 Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/832,883
; FILING DATE:
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 8321-13 US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 9837 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-832-883-68

Query Match 89.3%; Score 659.4; DB 1; Length 9837;
Best Local Similarity 95.6%; Pred. No. 9.9e-156;
Matches 689; Conservative 0; Mismatches 31; Indels 1; Gaps 1;
QY 18 GCGAATTCGAGCTCGGTACCCGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTG 77
DB 7217 GCACACTGGCGCGGTACTAGTGGATCCGAGCTCCGAGCTTGATGTCATAGCTTG 7158
QY 78 AGTATTCATAGTGTACCTAAATAGTGGCGTAATCATGTCATAGCTGTTTCCCTGTG 137
DB 7157 AGTATTCATAGTGTACCTAAATAGTGGCGTAATCATGTCATAGCTGTTTCCCTGTG 7098
QY 138 TGAATTTGTTATCGCTCACAAATTCACACAAATACGAGCGGGAAGCATAAAGTGTAAA 197
DB 7097 TGAATTTGTTATCGCTCACAAATTCACACAAATACGAGCGGGAAGCATAAAGTGTAAA 7038
QY 198 GCCTGGGTGCTTAATGAGTGAGCTAACTCACATTAATTCGTTGGCTCACTGCCCGCT 257
DB 7037 GCCT-GGGTGCCTAATGAGTGAGCTAACTCACATTAATTCGTTGGCTCACTGCCCGCT 6979
QY 258 TTCCAGTCCGGAAACCTGTCTGCGAGCTGCATTAAATGAATCGGCAACGCGCGGGAGA 317
DB 6978 TTCCAGTCCGGAAACCTGTCTGCGAGCTGCATTAAATGAATCGGCAACGCGCGGGAGA 6919
QY 318 GCGGTTTCGGTATTCGGCGCTCTTCGCTTCCTCGCTCACTACTCGCTCGGTC 377
DB 6918 GCGGTTTCGGTATTCGGCGCTCTTCGCTTCCTCGCTCACTACTCGCTCGGTC 6859
QY 378 GTTCGGCTCGGCGAGCGGTATTCAGCTCACTCAAGCGCGGTAAATAGCTTATCCAGAA 437
Best Local Similarity 95.6%; Pred. No. 9.9e-156;
Matches 689; Conservative 0; Mismatches 31; Indels 1; Gaps 1;
QY 18 GCGAATTCGAGCTCGGTACCCGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTG 77
DB 7217 GCACACTGGCGCGGTACTAGTGGATCCGAGCTCCGAGCTTGATGTCATAGCTTG 7158
QY 78 AGTATTCATAGTGTACCTAAATAGTGGCGTAATCATGTCATAGCTGTTTCCCTGTG 137
DB 7157 AGTATTCATAGTGTACCTAAATAGTGGCGTAATCATGTCATAGCTGTTTCCCTGTG 7098
QY 138 TGAATTTGTTATCGCTCACAAATTCACACAAATACGAGCGGGAAGCATAAAGTGTAAA 197
DB 7097 TGAATTTGTTATCGCTCACAAATTCACACAAATACGAGCGGGAAGCATAAAGTGTAAA 7038
QY 198 GCCTGGGTGCTTAATGAGTGAGCTAACTCACATTAATTCGTTGGCTCACTGCCCGCT 257
DB 7037 GCCT-GGGTGCCTAATGAGTGAGCTAACTCACATTAATTCGTTGGCTCACTGCCCGCT 6979
QY 258 TTCCAGTCCGGAAACCTGTCTGCGAGCTGCATTAAATGAATCGGCAACGCGCGGGAGA 317
DB 6978 TTCCAGTCCGGAAACCTGTCTGCGAGCTGCATTAAATGAATCGGCAACGCGCGGGAGA 6919
QY 318 GCGGTTTCGGTATTCGGCGCTCTTCGCTTCCTCGCTCACTACTCGCTCGGTC 377
DB 6918 GCGGTTTCGGTATTCGGCGCTCTTCGCTTCCTCGCTCACTACTCGCTCGGTC 6859
QY 378 GTTCGGCTCGGCGAGCGGTATTCAGCTCACTCAAGCGCGGTAAATAGCTTATCCAGAA 437
Best Local Similarity 95.6%; Pred. No. 9.9e-156;
Matches 689; Conservative 0; Mismatches 31; Indels 1; Gaps 1;
QY 18 GCGAATTCGAGCTCGGTACCCGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTG 77
DB 7217 GCACACTGGCGCGGTACTAGTGGATCCGAGCTCCGAGCTTGATGTCATAGCTTG 7158
QY 78 AGTATTCATAGTGTACCTAAATAGTGGCGTAATCATGTCATAGCTGTTTCCCTGTG 137
DB 7157 AGTATTCATAGTGTACCTAAATAGTGGCGTAATCATGTCATAGCTGTTTCCCTGTG 7098
QY 138 TGAATTTGTTATCGCTCACAAATTCACACAAATACGAGCGGGAAGCATAAAGTGTAAA 197
DB 7097 TGAATTTGTTATCGCTCACAAATTCACACAAATACGAGCGGGAAGCATAAAGTGTAAA 7038
QY 198 GCCTGGGTGCTTAATGAGTGAGCTAACTCACATTAATTCGTTGGCTCACTGCCCGCT 257
DB 7037 GCCT-GGGTGCCTAATGAGTGAGCTAACTCACATTAATTCGTTGGCTCACTGCCCGCT 6979
QY 258 TTCCAGTCCGGAAACCTGTCTGCGAGCTGCATTAAATGAATCGGCAACGCGCGGGAGA 317
DB 6978 TTCCAGTCCGGAAACCTGTCTGCGAGCTGCATTAAATGAATCGGCAACGCGCGGGAGA 6919
QY 318 GCGGTTTCGGTATTCGGCGCTCTTCGCTTCCTCGCTCACTACTCGCTCGGTC 377
DB 6918 GCGGTTTCGGTATTCGGCGCTCTTCGCTTCCTCGCTCACTACTCGCTCGGTC 6859
QY 378 GTTCGGCTCGGCGAGCGGTATTCAGCTCACTCAAGCGCGGTAAATAGCTTATCCAGAA 437

RESULT 15
US-08-832-877-68/c
Sequence 68, Application US/08832877
Patent No. 5840506
GENERAL INFORMATION:
APPLICANT: Giordano, Antonio
TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND PROGNOSIS OF
TITLE OF INVENTION: CANCER
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
STREET: Suite 1800 Two Penn Center Plaza
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/832,877
FILING DATE:
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 8321-13 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 9837 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-832-877-68

Query Match 89.3%; Score 659.4; DB 2; Length 9837;
Best Local Similarity 95.6%; Pred. No. 9.9e-156;
Matches 689; Conservative 0; Mismatches 31; Indels 1; Gaps 1;
QY 18 GCGAATTCGAGCTCGGTACCCGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTG 77
DB 7217 GCACACTGGCGCGGTACTAGTGGATCCGAGCTCCGAGCTTGATGTCATAGCTTG 7158
QY 78 AGTATTCATAGTGTACCTAAATAGTGGCGTAATCATGTCATAGCTGTTTCCCTGTG 137
DB 7157 AGTATTCATAGTGTACCTAAATAGTGGCGTAATCATGTCATAGCTGTTTCCCTGTG 7098
QY 138 TGAATTTGTTATCGCTCACAAATTCACACAAATACGAGCGGGAAGCATAAAGTGTAAA 197
DB 7097 TGAATTTGTTATCGCTCACAAATTCACACAAATACGAGCGGGAAGCATAAAGTGTAAA 7038
QY 198 GCCTGGGTGCTTAATGAGTGAGCTAACTCACATTAATTCGTTGGCTCACTGCCCGCT 257
DB 7037 GCCT-GGGTGCCTAATGAGTGAGCTAACTCACATTAATTCGTTGGCTCACTGCCCGCT 6979
QY 258 TTCCAGTCCGGAAACCTGTCTGCGAGCTGCATTAAATGAATCGGCAACGCGCGGGAGA 317
DB 6978 TTCCAGTCCGGAAACCTGTCTGCGAGCTGCATTAAATGAATCGGCAACGCGCGGGAGA 6919
QY 318 GCGGTTTCGGTATTCGGCGCTCTTCGCTTCCTCGCTCACTACTCGCTCGGTC 377
DB 6918 GCGGTTTCGGTATTCGGCGCTCTTCGCTTCCTCGCTCACTACTCGCTCGGTC 6859
QY 378 GTTCGGCTCGGCGAGCGGTATTCAGCTCACTCAAGCGCGGTAAATAGCTTATCCAGAA 437

```

Db 6858 GTTCGGCTGGCGGAGCGGTATCAGCTCACTCAAAGGCGGTAAATACGGTTATCCACAGAA 6799
QY 438 TCAGGGGATACGAGGAGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGT 497
Db 6798 TCAGGGGATACGAGGAGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGT 6739
QY 498 AAAAAGCCCGGTTGCTGGCGTTTTTTCATAGGCTCCGGCCCCCTGACGAGCATCACAAA 557
Db 6738 AAAAAGCCCGGTTGCTGGCGTTTTTTCATAGGCTCCGGCCCCCTGACGAGCATCACAAA 6679
QY 558 AATCGAGCTCAAGTCAGAGGTGGGAAACCCGACAGACTATAAAGTACCAGGCGTTT 617
Db 6678 AATCGAGCTCAAGTCAGAGGTGGGAAACCCGACAGACTATAAAGTACCAGGCGTTT 6619
QY 618 CCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCGGAGCTGCGGCTTACCGGATACCTG 677
Db 6618 CCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCGGAGCTGCGGCTTACCGGATACCTG 6559
QY 678 TCCGCCCTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCGTGTAGGTATCTC 737
Db 6558 TCCGCCCTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCGTGTAGGTATCTC 6499
QY 738 A 738
Db 6498 A 6498

```

Search completed: August 1, 2002, 09:31:23
Job time: 664 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2002, 08:39:47 ; Search time 2971.21 Seconds
(without alignments)
3352.423 Million cell updates/sec

Title: US-10-014-743-3

Perfect score: 738

Sequence: 1 ATACGACTCACTATAGGCG.....CTCAGCGTGTAGGTATCTCA 738

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estum:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	725.4	98.3	847	12	AZ687169 ENTMP05TF
2	713.4	96.7	838	12	AZ679307 ENTLS84TF
3	688.8	93.3	692	12	AQ040352 CIT-HSP-2
4	678.8	92.0	906	12	AZ549573 ENTFO03TF
5	665	90.1	666	12	AQ079096 CIT-HSP-2
6	661	89.6	891	12	AZ540838 ENTET16TF
7	653.8	88.6	683	10	BE662810 EST00554
8	652	88.3	652	12	AQ108633 CIT-HSP-2
9	651.4	88.3	653	12	AQ077504 CIT-HSP-2
10	651.2	88.2	667	10	BE430218 SUN000.D1
11	647.2	87.7	661	12	AQ057603 CIT-HSP-2
12	637.6	86.4	735	9	AV731085 AV731085
13	637	86.3	1039	9	AU081040 AU081040
14	637	86.3	1089	9	AU081124 AU081124
15	637	86.3	1163	9	AU081044 AU081044
16	636.6	86.3	1583	12	BH173601 BH173601
17	633.8	85.9	1067	9	AU081137 AU081137

18	631.4	85.6	633	12	AQ040787
19	627.6	85.0	642	12	AQ113213
20	625	84.7	996	9	AU081097
21	624.4	84.6	634	12	AQ038010
22	621.4	84.2	640	12	AQ074298
23	621.2	84.2	626	12	AQ11342
24	618.6	83.8	637	12	AQ077360
25	618.4	83.8	788	9	AJ281697
26	618	83.7	818	10	BI180988
27	615	83.3	623	12	AQ081172
28	609	82.5	1101	12	CNS00G5U
29	608.4	82.4	725	10	BI937939
30	606.2	82.1	735	10	BI938757
31	605.8	82.1	722	10	BI938061
32	605.6	82.1	623	10	BI937291
33	605.6	82.1	711	10	BI938755
34	604.4	81.9	686	10	BI941052
35	603.8	81.8	708	10	BI675834
36	603.4	81.8	605	12	AQ077353
37	599.8	81.3	689	10	BI937753
38	599.2	81.2	612	12	AQ080586
39	598.6	81.1	725	10	BI937927
40	597.6	81.0	704	10	BI938042
41	596.8	80.9	725	10	BI938759
42	594	80.5	704	10	BI937386
43	593.8	80.5	597	12	AQ040643
44	593.8	80.5	704	10	BI937399
45	590	79.9	663	10	BI940725

ALIGNMENTS

RESULT 1

AZ687169 847 bp DNA linear GSS 14-DEC-2000
ENTMP05TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
genomic, DNA sequence.
AZ687169
AZ687169.1 GI:11824315
GSS.
Entamoeba histolytica.
Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
1 (bases 1 to 847)
Lofthus,B., Van Aken,S. and Fraser,C.
Determination of clone end sequences from Entamoeba histolytica
HMI:IMSS sheared DNA library
Unpublished (2000)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjloftus@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 63
High quality sequence stop: 291.

FEATURES

source
1..847
/organism="Entamoeba histolytica"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHS01; Site1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a

method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999).

BASE COUNT 182 a 245 c 221 g 199 t

ORIGIN

Query Match 98.3%; Score 725.4; DB 12; Length 847;

Best Local Similarity 99.7%; Pred. No. 3.6e-199;

Matches 737; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
QY 1 ATAGGACTCTACTATAGG-GCGAATTCGAGCTCGGTACCGGGGATCCTCTAGAGTCGACC 59
Db 7 ATAGGACTCTACTATAGGTCGAATTCGAGCTCGGTACCGGGGATCCTCTAGAGTCGACC 66
QY 60 TGCAGGATCAAGCTTGTAGTATTCTATAGTGTACCTTAATAGCTTGGCGTAATCATGG 119
Db 67 TGCAGGATCAAGCTTGTAGTATTCTATAGTGTACCTTAATAGCTTGGCGTAATCATGG 126
QY 120 TCATAGCTGTTTCTGTGTGAATTTGTATCCGCTCAACATTCACACAAATAGGACC 179
Db 127 TCATAGCTGTTTCTGTGTGAATTTGTATCCGCTCAACATTCACACAAATAGGACC 186
QY 180 GGAACCATAAAGTGTAAAGCTGGGTCCTTAATGAGTGAGCTAACTACACATTAATTCG 239
Db 187 GGAACCATAAAGTGTAAAGCTGGGTCCTTAATGAGTGAGCTAACTACACATTAATTCG 246
QY 240 TTGGCTCACTGCCGCTTTCAGCTCGGGAACCTGTCTGCCAGCTGCATTAAATGAATC 299
Db 247 TTGGCTCACTGCCGCTTTCAGCTCGGGAACCTGTCTGCCAGCTGCATTAAATGAATC 306
QY 300 GGCCAAAGCGGGGAGAGCGGTTTCGTTATTTGGGCGCTCTCCGCTTCCTCGCTCACT 359
Db 307 GGCCAAAGCGGGGAGAGCGGTTTCGTTATTTGGGCGCTCTCCGCTTCCTCGCTCACT 366
QY 360 GACTCGCTCGCTCGCTCGCTCGGCGAGCGGTATCAGCTCACTCAAAAGCGGTA 419
Db 367 GACTCGCTCGCTCGCTCGCTCGGCGAGCGGTATCAGCTCACTCAAAAGCGGTA 426
QY 420 ATAGGTTATCCACAGAAATCAGGGGATACGCGAGGAAGAACATGTGAGCAAAAGGCCAG 479
Db 427 ATAGGTTATCCACAGAAATCAGGGGATACGCGAGGAAGAACATGTGAGCAAAAGGCCAG 486
QY 480 CAAAGGCCAGGACCGTAAAGGCGCGTGTGCTGGCGTTTTCATAGGCTCCGCCCC 539
Db 487 CAAAGGCCAGGACCGTAAAGGCGCGTGTGCTGGCGTTTTCATAGGCTCCGCCCC 546
QY 540 CCTGACGAGCATCACAAAATTCAGCGCTCAAGTCAGAGGTGGCGAAACCCACAGGACTA 599
Db 547 CCTGACGAGCATCACAAAATTCAGCGCTCAAGTCAGAGGTGGCGAAACCCACAGGACTA 606
QY 600 TAAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCGCTCTCCTGTTCCGACCCGT 659
Db 607 TAAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCGCTCTCCTGTTCCGACCCGT 666
QY 660 CCGCTTACCGGATACCTGTCGCGCTTTCCTTCGCGAAGCGTGGCGCTTCTCATAGC 719
Db 667 CCGCTTACCGGATACCTGTCGCGCTTTCCTTCGCGAAGCGTGGCGCTTCTCATAGC 726
QY 720 TCACGCTGTAGGTATCTCA 738
Db 727 TCACGCTGTAGGTATCTCA 745
```

RESULT 2

AZ679307

LOCUS

DEFINITION ENT1G847F Entamoeba histolytica 838 bp DNA linear GSS 14-DEC-2000

genomic, DNA sequence.

AZ679307

AZ679307.1 GI:11816453

GSS.

Entamoeba histolytica.

Entamoeba histolytica.

Eukaryota; Entamoebidae; Entamoeba.

1 (bases 1 to 838)

Loftus, B., Van Aken, S. and Fraser, C.

Determination of clone end sequences from Entamoeba histolytica

HMI:IMSS sheared DNA library

Unpublished (2000)

Contact: Brendan J Loftus

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0208

Fax: 301 838 3543

Email: bjoftus@tigr.org

Clones are derived from the Entamoeba histolytica HMI:IMSS sheared

DNA library

Seq primer: M13-Forward

Class: Shotgun

High quality sequence start: 52

High quality sequence stop: 297.

Location/Qualifiers

FEATURES

source

1..838

/organism="Entamoeba histolytica"

/strain="HMI:IMSS"

/db_xref="taxon:5759"

/clone_lib="Entamoeba histolytica sheared DNA"

/note="vector: pHOsl; Site 1: Bst I; Constructed at The

Institute for Genomic Research (TIGR), Rockville, MD.

Genomic DNA isolated from broth cultures of E. histolytica

using a method described by Clark and Diamond (Clark, a

C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a

method for isolate identification. Exp. Parasitol.

77:450.). The DNA was mechanically sheared to give a

tight size distribution (~2 kb). The v + i method used for

the library construction is described in detail in Smith,

H.O. and Venter, J.C. (Making small insert libraries for

whole genome shotgun sequencing projects. In Genome

Sequencing: A Practical Approach, eds. M. Vaudin and B.

Barell, Oxford University Press, 1999).

186 a 235 c 221 g 196 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 96.7%; Score 713.4; DB 12; Length 838;

Matches 717; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 16 GGGCAATTTCGAGCTCGGTACCGGGGATCCTCTAGAGTCGACCTGCGGATGCAAGCT 75

Db 29 GGTCAATTTCGAGCTCGGTACCGGGGATCCTCTAGAGTTTTCGAGGATGCAAGCT 88

QY 76 TGAGTATTCTATAGTGTACCTTAATAGCTTGGCGTAAATCATGCTAGCTGTTTCCTG 135

Db 89 TGAGTATTCTATAGTGTACCTTAATAGCTTGGCGTAAATCATGCTAGCTGTTTCCTG 148

QY 136 TGTGAAATTTGTTATCCGCTCACAAATTCACACAAATAGCAGCGGAGCAATAGTGA 195

Db 149 TGTGAAATTTGTTATCCGCTCACAAATTCACACAAATAGCAGCGGAGCAATAGTGA 208

QY 196 AAGCCTGGGTCCTTAATGAGTGAGCTAACTCACATTAATTCGTTGCCCTCACTGCCCG 255

Db 209 AAGCCTGGGTCCTTAATGAGTGAGCTAACTCACATTAATTCGTTGCCCTCACTGCCCG 268

QY 256 CTTTCCAGTCGGGAAACCTGTGTCGCCAGCTGCATTAATTAATCGGCCAACCGCGGGA 315

Db 269 CTTTCCAGTCGGGAAACCTGTGTCGCCAGCTGCATTAATTAATCGGCCAACCGCGGGA 328

QY 316 GAGGGGTTTGGGATTTGGGCGCTCTTCGGCTTCCTCGCTCACTGCTGCGCTCGG 375

```

Db 329 GAGCGGTTTTCGCTATTGGCGCTCTTCCGCTTCTCGCTCACTCGCTCGCTCGG 388
QY 376 TCCTTCGGCTCGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATCCACAG 435
Db 389 TCCTTCGGCTCGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATCCACAG 448
QY 436 AATCAGGGGTAACGCGAGGAAAGAAATGTCAGCAAAAGGCGGCAAAAGGCGGAGAAC 495
Db 449 AATCAGGGGTAACGCGAGGAAAGAAATGTCAGCAAAAGGCGGCAAAAGGCGGAGAAC 508
QY 496 GTAAAAAGCGCGGTTCTCGCTGCGCTTTTTCATAGGCTCCGCCCCCTTACGAGCATCACA 555
Db 509 GTAAAAAGCGCGGTTCTCGCTGCGCTTTTTCATAGGCTCCGCCCCCTTACGAGCATCACA 568
QY 556 AATATCAGCGCTCAAGTCAGAGGTGGCGAAACCGGACGAGCATATAAGATACAGGGGT 615
Db 569 AATATCAGCGCTCAAGTCAGAGGTGGCGAAACCGGACGAGCATATAAGATACAGGGGT 628
QY 616 TTCCCGCTGGAAGCTCCCTCGTGGGCTCTCTGTTCCGACCGCTCCGCTTACCGGATAC 675
Db 629 TTCCCGCTGGAAGCTCCCTCGTGGGCTCTCTGTTCCGACCGCTCCGCTTACCGGATAC 688
QY 676 TGTCCGCTTTCTCCCTTCCGGAAGCGGTGGCGCTTTCTCATAGCTCAGCGCTGTAGGTATC 735
Db 689 TGTCCGCTTTCTCCCTTCCGGAAGCGGTGGCGCTCTCTCATAGCTCAGCGCTGTAGGTATC 748
QY 736 TCA 738
Db 749 TCA 751

```

RESULT 3

AQ040352 692 bp DNA linear GSS 11-JUL-1998
 CIT-HSP-2327K21.TF CIT-HSP Homo sapiens genomic clone 2327K21, DNA sequence.

ACCESSION AQ040352

VERSION AQ040352.1 GI:3306184

KEYWORDS GSS.

SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 692)
 Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
 Golden,K., Berry,K., Grainger,D., Suh,E., Wible,C., Shizuya,H.,
 Simon,M., and Venter,J.C.
 Use of a random BAC End Sequence Database for Sequence-Ready Map
 Building (1998)

JOURNAL

Unpublished (1998)
 Other_GSSs: CIT-HSP-2327K21.TR
 Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadams@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: M13-21
 Class: BAC ends.

FEATURES

source

Location/Qualifiers
 1..692
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="2327K21"
 /clone_lib="CIT-HSP"
 /sex="Male"
 /cell_type="Sperm"
 /note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:

```

BASE COUNT 158 a 193 c 186 g 155 t
ORIGIN
Query Match 93.3%; Score 688.8; DB 12; Length 692;
Best Local Similarity 99.7%; Pred. No. 1.3e-188;
Matches 690; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 16 GGGCAATTCGAGCTCGGTACCCGGGATCCTCTAGAGTCGACCTCAGGATCGCAAGCT 75
Db 1 GGGCAATTCGAGCTCGGTACCCGGGATCCTCTAGAGTCGACCTCAGGATCGCAAGCT 60
QY 76 TGAGTATTCTATAGTGTCACTAAATAGCTTGGCGTAAATCATGTGCTATGCTTTCCCTG 135
Db 61 TGAGTATTCTATAGTGTCACTAAATAGCTTGGCGTAAATCATGTGCTATGCTTTCCCTG 120
QY 136 TGTGAAATTTGTTATCCGCTCACAAATCCACAAACATACGAGCGGGAAGCATAAAGTGTA 195
Db 121 TGTGAAATTTGTTATCCGCTCACAAATCCACAAACATACGAGCGGGAAGCATAAAGTGTA 180
QY 196 AAGCCTGGGTCCTTAATGAGTGAGCTAACTACATTAATTTGGTTGGGCTCACTGCGCG 255
Db 181 AAGCCTGGGTCCTTAATGAGTGAGCTAACTACATTAATTTGGTTGGGCTCACTGCGCG 240
QY 256 CTTTCCAGTCGGAACCTGTGTCGCCAGCTGATTAATCAATCGGCCACGCGCGGGGA 315
Db 241 CTTTCCAGTCGGAACCTGTGTCGCCAGCTGATTAATCAATCGGCCACGCGCGGGGA 300
QY 316 GAGCGGTTTGGCTATTGGGCGCTCTTCCGCTTCCCTCGCTCACTGACTCGCTCGCTCGG 375
Db 301 GAGCGGTTTGGCTATTGGGCGCTCTTCCGCTTCCCTCGCTCACTGACTCGCTCGG 360
QY 376 TCGTTTCGGTCGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATCCACAG 435
Db 361 TCGTTTCGGTCGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATCCACAG 420
QY 436 AATCAGGGGTAACCGAGGAAACATGTGAGCAAAAGGCGGCAAAAGGCGGAGAAC 495
Db 421 AATCAGGGGTAACCGAGGAAACATGTGAGCAAAAGGCGGCAAAAGGCGGAGAAC 480
QY 496 GTAAAAAGCGCGGTTCTCGCTGCGCTTTTTCATAGGCTCCGCCCCCTTACGAGCATCACA 555
Db 481 GTAAAAAGCGCGGTTCTCGCTGCGCTTTTTCATAGGCTCCGCCCCCTTACGAGCATCACA 540
QY 556 AAAATCGAGCTCAAGTCAGAGGTGGCGAAACCGGACGAGCATATAAGATACCGAGCGT 615
Db 541 AAAATCGAGCTCAAGTCAGAGGTGGCGAAACCGGACGAGCATATAAGATACCGAGCGT 600
QY 616 TTCCCGCTGGAAGCTCCCTCGTGGGCTCTCTGTTCCGACCGCTTACCGGATAC 675
Db 601 TTCCCGCTGGAAGCTCCCTCGTGGGCTCTCTGTTCCGACCGCTTACCGGATAC 660
QY 676 TGTCCGCTTTCTCCCTTCCGGAAGCGGTGGCG 707
Db 661 TGTCCGCTTTCTCCCTTCCGGAAGCGGTGGCG 692

```

RESULT 4

AZ549573 906 bp DNA linear GSS 14-NOV-2000
 LOCUS ENTUF03TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
 DEFINITION genomic, DNA sequence.

ACCESSION AZ549573

VERSION AZ549573.1 GI:11174314

KEYWORDS GSS.

SOURCE Entamoeba histolytica.

ORGANISM Entamoeba histolytica.

REFERENCE 1 (bases 1 to 906)

Loftus,B., Van Aken,S. and Fraser,C.

Determination of clone end sequences from Entamoeba histolytica

HMI:IMSS sheared DNA library

JOURNAL
COMMENT

Unpublished (2000)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjlouftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 141
High quality sequence stop: 293.
Location/Qualifiers
1. .906
/organism="Entamoeba histolytica"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/note="Vector: phOS1; Site 1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, I.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."

BASE COUNT
ORIGIN

222 a 247 c 230 g 207 t
Query Match 92.0%; Score 678.8; DB 12; Length 906;
Best Local Similarity 95.0%; Pred. No. 1.2e-185;
Matches 701; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 ATAGGACTACTATAGGGCAATTCGAGCTCGGTACCGGGGATCCTCTAGATCGACCT 60
DB 10 ATACGACTGACTATAGGGCAATTCGAGCTCGGTACCGGGGATCCTCTAGATCGACCT 69
QY 61 GCAGGACTCAAGCTTGAGTATCTATAGTGTACCTAAATAGCTTGGCGTAATCATGGT 120
DB 70 GCATGATGCAAGCTTGAGTATCTATAGTGTACCTAAATAGCTTGGCGTAATCATGGT 129
QY 121 CATAGCTGTTTCTGTTGTAATTTGTTTATCCGCTCACAATTCACACATACGAGCGG 180
DB 130 CATAGCTGTTTCTGTTGTAATTTGTTTATCCGCTCACAATTCACACATACGAGCGG 189
QY 181 GAAGCATAAAGTGAAGCCCTGGGTGCTTAATGAGTGAGCTACTACTAATTAATTCGGT 240
DB 190 GAAGCATAAAGTGAAGCCCTGGGTGCTTAATGAGTGAGCTACTACTAATTAATTCGGT 249
QY 241 TGGCGCTCACTGCCCGCTTTCAGTCGCGGAACCTGTGTCGACAGTCGATTAATGAATCG 300
DB 250 GCGCGCTCACTGCCCGCTTTCAGTCGCGGAACCTGTGTCGACAGTCGATTAATGAATCG 309
QY 301 GCCAACGCGGGGAGAGCGGGTTTGGGTATTTGGCGCTCTTCCGCTTCCTCGCTCACTG 360
DB 310 GCCAACGCGGGGAGAGCGGGTATTTGGGTATTTGGCGCTCTTCCGCTTCCTCGCTCACTG 369
QY 361 ACTCGCTGCGCTCGGTGTTGCGCTGGCGGAGCGGGTATGAGTCACTCAACAGCGGTAA 420
DB 370 ACTCGCTGCGCTCGGTGTTGCGCTGGCGGAGCGGGTATGAGTCACTCAACAGCGGTAA 429
QY 421 TAGCGTTATCCAGAAATCAGGGGATTAACGAGGAAAGAACATCTGAGCAAAAGGCCAGC 480
DB 430 TAGCGTTATCCAGAAATCAGGGGATTAACGAGGAAAGAACATCTGAGCAAAAGGCCAGC 489

QY 481 AAAAGCCAGGAACCGTAAAAAGCCGCTGCTGGCGTCTTTTCCATAGGCTCCGCCGCC 540
DB 490 ATAATGCCATGAACCGTAAAAAGCCGCTGCTGGCGTCTTTTCCATAGGCTCCGCCGCC 549
QY 541 CTGAGGACATCACAAAAATCGACGCTCAAGTCAGAGTGGGGAACCCGACGAGCTAT 600
DB 550 CTGAGGACATCACAAAAATCGACGCTCAAGTCAGAGTGGGGAACCCGACGAGCTAT 609
QY 601 AAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTGCGGCTCTCTGTTCCGACCCCTGC 660
DB 610 TAAGATACCATGCGTTTCCCGCTGGAAGCTCAAGCGCTCTCTGTTACGACCCCTGC 669
QY 661 CGTTTACCGGATACCTGCGGCTTTCTCCCTTCGGGAGCGCTGGCGCTTCTCATAGCT 720
DB 670 CGTTTACCGGATACCTGCGGCTTTCTACCTTTCAGCGAGCGCTGGCGCTTAACCTATAGCT 729
QY 721 CACGCTGAGGTATCTCA 738
DB 730 CACGCTGATGTATCTCA 747

RESULT 5

LOCUS AQ079096 666 bp DNA linear GSS 20-AUG-1998
DEFINITION CIT-HSP-2356E24.TF CIT-HSP Homo sapiens genomic clone 2356E24, DNA
sequence.
ACCESSION AQ079096
VERSION AQ079096.1 GI:3440280
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 666)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/hungen/tldb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES

source Location/Qualifiers
1. .666
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2356E24"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT 156 a 186 c 174 g 149 t 1 others
ORIGIN

Query Match 90.1%; Score 665; DB 12; Length 666;
Best Local Similarity 99.8%; Pred. No. 1e-181;
Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 AATTCGAGCTCGGTACCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGAGT 80
DB 1 AATTCGAGCTCGGTACCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGAGT 60

Qy	81	ATTCTATAGTGTACCTAAATAGCTTGGCGTAAATCATAGTATAGCTGTTTCTCTGTGTGA	140
Db	61	ATTCTATAGTGTACCTAAATAGCTTGGCGTAAATCATAGTATAGCTGTTTCTCTGTGTGA	120
Qy	141	AATTGTTATCCGCTCAAAATTCACACAAACATACGAGCCGGAAGCATAAAGTGTAAAGCC	200
Db	121	AATTGTTATCCGCTCAAAATTCACACAAACATACGAGCCGGAAGCATAAAGTGTAAAGCC	180
Qy	201	TGGGTTGCCCTAATGAGTGAAGCTAACTACATTAATTGCGTTGGCTTACCTGCCCGTTTC	260
Db	181	TGGGTTGCCCTAATGAGTGAAGCTAACTACATTAATTGCGTTGGCTTACCTGCCCGTTTC	240
Qy	261	CAGTCGGGAAACCTGCTGCCAGCTGCATTAATGAATCGGCCAAGCGCGGGGAGAGGC	320
Db	241	CAGTCGGGAAACCTGCTGCCAGCTGCATTAATGAATCGGCCAAGCGCGGGGAGAGGC	300
Qy	321	GGTTTGGGTAAATGGGCGCTCTTCGGTTCCTCGCTCACTGACTCGTGGCTCGGTCGTT	380
Db	301	GGTTTGGGTAAATGGGCGCTCTTCGGTTCCTCGCTCACTGACTCGTGGCTCGGTCGTT	360
Qy	381	CGCTTCGGCGAGCGGTATCAGCTCACTCAAAGCGGTAATACGGTTATCCACAGAATCA	440
Db	361	CGCTTCGGCGAGCGGTATCAGCTCACTCAAAGCGGTAATACGGTTATCCACAGAATCA	420
Qy	441	GGGGATAACGACGAAAGAACATGTGAGCAAAAGGCCACAAAGGCCAGGAACCGTAAA	500
Db	421	GGGGATAACGACGAAAGAACATGTGAGCAAAAGGCCACAAAGGCCAGGAACCGTAAA	480
Qy	501	AAGSCCGGTGCTGGCGTTTTTCCATAGCTCCGCCCCCTGACGAGATCACAAAAT	560
Db	481	AAGSCCGGTGCTGGCGTTTTTCCATAGCTCCGCCCCCTGACGAGATCACAAAAT	540
Qy	561	CGACGCTCAAGTCAGAGTGGGAAACCCGACAGGACTATAAGATACACAGCGTTTCC	620
Db	541	CGACGCTCAAGTCAGAGTGGGAAACCCGACAGGACTATAAGATACACAGCGTTTCC	600
Qy	621	CTTGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCC	680
Db	601	CTTGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCC	660
Qy	681	GCCTTT	686
Db	661	GNCTTT	666

RESULT	6	
LOCUS	AZ540838	
DEFINITION	ENT16TF Entamoeba histolytica Sheared DNA sequence.	
ACCESSION	AZ540838	
VERSION	AZ540838.1	
KEYWORDS	GI:11148010	
SOURCE	GSS.	
ORGANISM	Entamoeba histolytica.	
REFERENCE	Entamoeba histolytica.	
AUTHORS	Eukaryota; Entamoebidae; Entamoeba.	
TITLE	1 (bases 1 to 891)	
JOURNAL	Loftus, B., Van Aken, S. and Fraser, C.	
COMMENT	Determination of clone end sequences from Entamoeba histolytica HM1:IMSS sheared DNA library Unpublished (2000) Contact: Brendan J Loftus Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0208 Fax: 301 838 3543 Email: b.loftus@tigr.org Clones are derived from the Entamoeba histolytica HM1:IMSS sheared DNA library Seq primer: M13-Forward Class: shotgun	

```

High quality sequence start: 68
High quality sequence stop: 282.
Location/Qualifiers
1. .891
/organism="Entamoeba histolytica"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone.lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOS1; Site_1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making Small Insert Libraries for
Whole Genome Shotgun Sequencing Projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barelli, Oxford University Press, 1999)."
```

BASE COUNT	199 a	255 c	213 g	224 t
ORIGIN	Bureau of Oxford University Press, 1955).			
Query Match	89.6%			
Best Local Similarity	96.3%			
Matches	709; Conservative	0; Mismatches	25; Indels	3; Gaps
				3;

[illegible]

QY 663 CTTACCGGATACCTGTCGGCGCTTCTCCCTTCGGGAAGCGTGG-CGCTTCTCATAGCTC 721
 Db 661 CTTACCGGATACCTGTCGGCGCTTCTCCCTTCGGGAAGCGTGGACGCTTCTCATAGCTC 720

QY 722 ACCTGTAGGTATCTCA 738
 Db 721 ACCTGGAAGGTTCTCA 737

RESULT 7
 BE662810
 LOCUS BE662810 683 bp mRNA linear EST 01-JUL-2001
 DEFINITION EST00554 Arabidopsis Chronic Ozone Forward-Subtracted Library
 Arabidopsis thaliana cDNA clone AtCOZ1G4, mRNA sequence.
 ACCESSION BE662810
 VERSION BE662810.1 GI:14580073
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 683)
 Eckardt,N.A., Higgins,B.K., Schloss,A. and Fedoroff,N.V.
 Subtractive cloning and microarray expression analysis of
 ozone-induced genes in Arabidopsis
 Unpublished (2000)
 Contact: Nina V. Fedoroff
 Biotechnology Institute
 Penn State University
 519 Wartik Lab, University Park, PA 16802, USA
 Tel: 8148635717
 Fax: 8148631357
 Email: nvf@psu.edu
 Sequence does not include adaptor sequences (corresponding to
 Clontech PCR-Select adaptors 1 and 2R) present on either side of
 insert.
 Insert Length: 794 Std Error: 200.00
 Seq primer: Clontech PCR-Select Nested Primer 2R
 High quality sequence start: 26
 High quality sequence stop: 708
 POLYA-No. Location/Qualifiers
 1. 683
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /clone="AtCoz1G4"
 /clone_lib="Arabidopsis Chronic Ozone Forward-Subtracted
 Library"
 /tissue_type="Leaf"
 /note="Vector: pT-Adv; PCR Suppression Subtractive
 Hybridization library (Clontech PCR-Select). 'Tester' =
 RNA pooled from leaf tissue of plants exposed to 3, 6,
 and 9 days of ozone at 0.15 ul/L for 6 h/d. 'Driver' =
 leaf tissue from control plants maintained in clean air.
 Cloned into 3.9 kb AdvantaAge vector (Clontech)."
 157 a 193 c 176 g 155 t 2 others

BASE COUNT 157 a 193 c 176 g 155 t 2 others
 ORIGIN

Query Match 88.6%; Score 653.8; DB 10; Length 683;
 Best Local Similarity 98.5%; Pred. No. 1.9e-178;
 Matches 669; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 16 GGCGAATTCGAGTCGGTACCGGGGATCCTCTAGAGTCGACCTCGAGGATCGAAGCT 75
 Db 4 GGCGAATTCGAGTCGGTACCGGGGATCCTCTAGAGTCGACCTCGAGGATCGAAGCT 63

QY 76 TGAGTATTCATAGTGTACCACTAATAGCTTGGCGTAATCATGTGTCATAGCTGTTTCCTG 135
 Db 64 TGAGTATTCATAGTGTACCACTAATAGCTTGGCGTAATCATGTGTCATAGCTGTTTCCTG 123

RESULT 8
 AQ108633
 LOCUS AQ108633 652 bp DNA linear GSS 29-AUG-1998
 DEFINITION CIT-HSP-2379M3.TF CIT-HSP Homo sapiens genomic clone 2379M3, DNA
 sequence.
 ACCESSION AQ108633
 VERSION AQ108633.1 GI:3485323
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 652)
 Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
 Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
 Venter,J.C.
 Use of a random human BAC End Sequence Database for Sequence-Ready
 Map Building
 Unpublished (1998)
 Other_GSSs: CIT-HSP-2379M3.TR
 Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: madams@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
 http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: M13-21
 Class: BAC ends.

```
FEATURES
source
Location/Qualifiers
1..652
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2379M3"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2: HindIII"
HindIII
152 a 184 c 170 g 146 t

Query Match 88.3%; Score 652; DB 12; Length 652;
Best Local Similarity 100.0%; Pred. No. 5.9e-178;
Matches 652; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 GGGGATCCCTCTAGAGTCGACCTCGAGGCATGCAAGCTTGAGTATTTCTATAGTGTACGCTA 98
Db 1 GGGGATCCCTCTAGAGTCGACCTCGAGGCATGCAAGCTTGAGTATTTCTATAGTGTACGCTA 60

Qy 99 AATAGTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATGTTATCGCTCACA 158
Db 61 AATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATGTTATCGCTCACA 120

Qy 159 ATTCCACACAACATAGAGCGGCAAGCATAAAGCTGTAAGCCCTGGGTGCTAATGAGTG 218
Db 121 ATTCCACACAACATAGAGCGGCAAGCATAAAGCTGTAAGCCCTGGGTGCTAATGAGTG 180

Qy 219 AGTAACTCACATTAATTCGGTTGCGCTCAGTCCCGCTTTCCAGTCGGGAAACCTGTGC 278
Db 181 AGTAACTCACATTAATTCGGTTGCGCTCAGTCCCGCTTTCCAGTCGGGAAACCTGTGC 240

Qy 279 TGGCAGCTGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTGCGTATTTGGCGC 338
Db 241 TGGCAGCTGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTGCGTATTTGGCGC 300

Qy 339 TCTTCGCGTTCCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGCGAGCGGTA 398
Db 301 TCTTCGCGTTCCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGCGAGCGGTA 360

Qy 399 TCAGCTCACTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGCGAGGAAG 458
Db 361 TCAGCTCACTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGCGAGGAAG 420

Qy 459 ACATGTGAGCAAAAGCGCAAGAAAGCGGAGGACCGTAAAAGCGCGGTTGCTGGCG 518
Db 421 AACATGTGAGCAAAAGCGCAAGAAAGCGGAGGACCGTAAAAGCGCGGTTGCTGGCG 480

Qy 519 TTTTTCATAGGCTCGCGCCCTCGAGCAGCATCAAAAATCGAGCTCAAGTCAGAGG 578
Db 481 TTTTTCATAGGCTCGCGCCCTCGAGCAGCATCAAAAATCGAGCTCAAGTCAGAGG 540

Qy 579 TGGCGAAACCCGAGGACTATAAAGATACAGCGCTTTCCCGCTGGAAGCTCCCTCGTG 638
Db 541 TGGCGAAACCCGAGGACTATAAAGATACAGCGCTTTCCCGCTGGAAGCTCCCTCGTG 600

Qy 639 CGCTCTCTGTTCCGACCTTCCGCTTACCGGATACCTGTCGCGCTTTCTCC 690
Db 601 CGCTCTCTGTTCCGACCTTCCGCTTACCGGATACCTGTCGCGCTTTCTCC 652

RESULT 9
AQ077504
LOCUS
DEFINITION
CIT-HSP-2365K21.TF CIT-HSP Homo sapiens genomic clone 2365K21, DNA
sequence.
ACCESSION
AQ077504
VERSION
AQ077504.1
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
```

```
REFERENCE
AUTHORS
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL
Unpublished (1998)
COMMENT
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES
source
Location/Qualifiers
1..653
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2365K21"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2: HindIII"
HindIII
158 a 177 c 173 g 145 t

Query Match 88.3%; Score 651.4; DB 12; Length 653;
Best Local Similarity 99.8%; Pred. No. 8.8e-178;
Matches 652; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATAGCACTCACTATAGGCGGAATTCGAGCTCGGTACCGGGGATCCTCTAGAGTCGACCT 60
Db 1 ATAGCACTCACTATAGGCGGAATTCGAGCTCGGTACCGGGGATCCTCTAGAGTCGACCT 60

Qy 61 CGAGCATGCAAGCTTGAGTATTCATAGTGTACCTAAATAGCTGGCGTAAATCATGT 120
Db 61 CGAGCATGCAAGCTTGAGTATTCATAGTGTACCTAAATAGCTGGCGTAAATCATGT 120

Qy 121 CATAGCTGTTTCTGTGTGAAATTTATTCGCTCACAATTCACACACATACGAGCG 180
Db 121 CATAGCTGTTTCTGTGTGAAATTTATTCGCTCACAATTCACACACATACGAGCG 180

Qy 181 GAAGCATAAAGTAAAGCCCTGGGTGCTTATGAGTGAAGTCACTCACTTAATTTGCT 240
Db 181 GAAGCATAAAGTAAAGCCCTGGGTGCTTATGAGTGAAGTCACTCACTTAATTTGCT 240

Qy 241 TGGCGTCACTCGCGCTTTCCAGTCGGGAAACCTGCTGTCGAGCTGCACTTAATGAATCG 300
Db 241 TGGCGTCACTCGCGCTTTCCAGTCGGGAAACCTGCTGTCGAGCTGCACTTAATGAATCG 300

Qy 301 GCCAACGCGCGGAGAGCGGTTTTCGATTTGGGCGTCTTTCGCTTCCTCGCTACATG 360
Db 301 GCCAACGCGCGGAGAGCGGTTTTCGATTTGGGCGTCTTTCGCTTCCTCGCTACATG 360

Qy 361 ACTCGCTCGGCTCGGCTGCTGCGGAGCGGCTATCAGCTCACTCAAAAGCGGTA 420
Db 361 ACTCGCTCGGCTCGGCTGCTGCGGAGCGGCTATCAGCTCACTCAAAAGCGGTA 420

Qy 421 TACGTTATCCACAGAAATCAGGGGATAACGAGGAAAGAAATGTCAGCAAAAGGCCAGC 480
Db 421 TACGTTATCCACAGAAATCAGGGGATAACGAGGAAAGAAATGTCAGCAAAAGGCCAGC 480

Qy 481 AAAAGGCGAGGAACCGTAAAAGCGCGGTTGCTGCGCTTTTCCATAGGCTCCGCC 540
Db 481 AAAAGGCGAGGAACCGTAAAAGCGCGGTTGCTGCGCTTTTCCATAGGCTCCGCC 540
```

Db 481 AAGAGCCAGAACCGTAAAGAGCCGCGTGTGCTGGCGTTTTTCCATAGAGCTCCGTCGCC 540
QY 541 CTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACCCGACAGACTAT 600
Db 541 CTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACCCGACAGACTAT 600
QY 601 AAGATACCGAGCGTTTCCCGCTGGAGCTCCCTCGTGGCGCTCCTGTTCCG 653
Db 601 AAGATACCGAGCGTTTCCCGCTGGAGCTCCCTCGTGGCGCTCCTGTTCCG 653

RESULT 10
BE430218
LOCUS
DEFINITION
ACCESSTION
VERSION
KEYWORDS
SOURCE
ORGANISM

BE430218 667 bp mRNA linear EST 24-JUL-2000
SUN000.D12R990802 ITEC SUN Wheat cDNA Library Triticum aestivum
cDNA clone SUN000.D12, mRNA sequence.
BE430218
BE430218.1 GI:9428061
EST.
bread wheat.
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.

REFERENCE
AUTHORS
JOURNAL
COMMENT

Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier,
'S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P.,
Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J., Joudrier, P.,
Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P., Ogihara, Y.,
Pecchioni, N., Quilset, C., Schuch, W., Selvaraj, G., Shariflou, M.,
Sorells, M., Warburton, M. and Wenzel, G.
International Triticeae EST Cooperative (ITEC): Production of
Expressed Sequence Tags for Species of the Triticeae
Unpublished (2000)
Contact: Shariflou M
Plant Breeding Institute, University of Sydney
107 Cobbitty Rd., Cobbitty NSW 2570 Australia
Tel: 61 2 9351 8803
Fax: 61 2 9351 8851
Email: msharif@plop.usyd.edu.au
International Triticeae EST Cooperative (ITEC)
http://wheat.pw.usda.gov/genome.

FEATURES
source
Location/Qualifiers
1..667
/organism="Triticum aestivum"
/db_xref="taxon:4565"
/clone="SUN000.D12"
/clone_lib="ITEC SUN Wheat cDNA Library"

BASE COUNT 152 a 190 c 175 g 150 t
ORIGIN

Query Match 88.2%; Score 651.2; DB 10; Length 667;
Best Local Similarity 99.4%; Pred. No. 1e-177; 3; Indels 1; Gaps 1;
Matches 664; Conservative 0; Mismatches 1; Gaps 1;

QY 29 CTCGGTACCGGGGATCCTCTAGAGTGGACCTGCAGGCATGCAAGCTTGAGTATCTATA 88
Db 1 CTCGGTACCGGGGATCCTCTAGAGTGGACCTGCAGGCATGCAAGCTTGAGTATCTATA 60
QY 89 GTCTACCTAAATAGCTTGGCGTAATCATGTCATAGCTGTTTCCCTGTGCAATTTGTTA 148
Db 61 GTCTACCTAAATAGCTTGGCGTAATCATGTCATAGCTGTTTCCCTGTGCAATTTGTTA 120
QY 149 TCCGCTCACAAATCCACACAACTACGAGCGCGGAAGCATAAAGTGAAGCCCTGGCGTGC 208
Db 121 TCCGCTCACAAATCCACACAACTACGAGCGCGGAAGCATAAAGTGAAGCCCTGGCGTGC 180
QY 209 CTAATGAGTGAGTAACTCACTAAATTTGGCTTGGCTCACTGCCCGCTTTCCAGTCGGG 268
Db 181 CTAATGAGTGAGTAACTCACTAAATTTGGCTTGGCTCACTGCCCGCTTTCCAGTCGGG 240
QY 269 AACCTGTGCTGCCAGCTGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTGGC 328

Db 241 AAACCTGTGCTGCCAGCTGCATTAATGAATCGCCCAACGCGGGGAGAGCGGTTTGGC 300
QY 329 TATTGGCGCGCTCTTCCGCTTCTCCTCCTCACTCACTCGCTCGCTCGCTCGTTCGGCTGGC 388
Db 301 TATTGGCGCGCTCTTCCGCTTCTCCTCCTCACTCACTCGCTCGCTCGCTCGTTCGGCTGGC 360
QY 389 GCGAGCGGTATCAGCTCACTCAAGCGCGTAATACGGTTTATCCACAGAATCAGGGGATAA 448
Db 361 GCGAGCGGTATCAGCTCACTCAAGCGCGTAATACGGTTTATCCACAGAATCAGGGGATAA 420
QY 449 GCGAGGAAAGACATGTGAGCAAAAGGCGCAGCAAAAGGCGCAGGAAACCGTAAAGGCGCGC 508
Db 421 GCGAGGAAAG-ACATGTGAGCAAAAGGCGCAGCAAAAGGCGCAGGAAACCGTAAAGGCGCGC 479
QY 509 GTTGTGCGCGTTTTCATAGGCTCGCCCGCTCAGCAGCATCAAAAGGCGCAGGCGTC 568
Db 480 GTTGTGCGCGTTTTCATAGGCTCGCCCGCTCAGCAGCATCAAAAGGCGCAGGCGTC 539
QY 569 AAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAAGGCTTTCCCGCTGGAAG 628
Db 540 AAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAAGGCTTTCCCGCTGGAAG 599
QY 629 CTCCTCTGCGCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCT 688
Db 600 CTCCTCTGCGCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCT 659
QY 689 CCCCTCGG 696
Db 660 CTTTCGGG 667

RESULT 11
A0057603
LOCUS
DEFINITION
ACCESSTION
VERSION
KEYWORDS
SOURCE
ORGANISM

A0057603 661 bp DNA linear GSS 30-JUL-1998
CIT-HSP-2345M2.TF CIT-HSP Homo sapiens genomic clone 2345M2, DNA
sequence.
A0057603
A0057603.1 GI:3354129
GSS.
human.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 661)
Adams, M.D., Rounsley, S.D., Zhao, S., Field, C.E., Bass, S., Linher, K.,
Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H.,
Simon, M. and Venter, J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
Unpublished (1998)
Other GSSs: CIT-HSP-2345M2.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamads@tigr.org
Clones are available from Research Genetics (Info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES
source
Location/Qualifiers
1..661
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2345M2"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:

BASE COUNT 159 a 177 c 175 g 150 t
 HINDIII*

Query Match 87.7%; Score 647.2; DB 12; Length 661;
 Best Local Similarity 99.5%; Pred. No. 1.5e-176;
 Matches 649; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 TAGAGTCTACATAGGCGAATTCGAGCTCGGTACCGGGATCCTCTAGAGTCGAGCTG 61
 DB 10 TAGAGTCTACATAGGCGAATTCGAGCTCGGTACCGGGATCCTCTAGAGTCGAGCTG 69
 QY 62 CAGGCGATCAAGCTTGAGTATTTCTATAGTCTCACTAAATAGCTTGGCGTAATCATGGTC 121
 DB 70 CAGGCGATCAAGCTTGAGTATTTCTATAGTCTCACTAAATAGCTTGGCGTAATCATGGTC 129
 QY 122 ATAGCTGTTTCCCTGTGTGAATTTATTCGGCTCACAAATTCACACACATACGAGCGG 181
 DB 130 ATAGCTGTTTCCCTGTGTGAATTTATTCGGCTCACAAATTCACACACATACGAGCGG 189
 QY 182 AAGCATAAAGTGTAAAGCTTGGGTGCTAATGAGTACGTAAGTCAATTAATGGCTT 241
 DB 190 AAGCATAAAGTGTAAAGCTTGGGTGCTAATGAGTACGTAAGTCAATTAATGGCTT 249
 QY 242 GCCTCTCACTGCGCGCTTTCCAGTCGGGAAACCTGCTGCGAGCTGCATTAATGAATCGG 301
 DB 250 GCCTCTCACTGCGCGCTTTCCAGTCGGGAAACCTGCTGCGAGCTGCATTAATGAATCGG 309
 QY 302 CCAACGCGGGGAGAGCGGTTTGGCTATTGGGCGCTCTTCGCTTTCCTGGCTCACTGA 361
 DB 310 CCAACGCGGGGAGAGCGGTTTGGCTATTGGGCGCTCTTCGCTTTCCTGGCTCACTGA 369
 QY 362 CTGCTGCGGCTCGGCTTGGCTTGGCTGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAAT 421
 DB 370 CTGCTGCGGCTCGGCTTGGCTTGGCTGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAAT 429
 QY 422 ACGTTATCCACAGATTCAGGGATAACGAGGAAAGAAAGTGTGAGCAAAAGGCCAGCA 481
 DB 430 ACGTTATCCACAGATTCAGGGATAACGAGGAAAGAAAGTGTGAGCAAAAGGCCAGCA 489
 QY 482 AAAGGCCAGAACCGTAAAAAGCGCGTTGCTGGCGGTTTTTCATAGGCTTCCGCCGCC 541
 DB 490 AAAGGCCAGAACCGTAAAAAGCGCGTTGCTGGCGGTTTTTCATAGGCTTCCGCCGCC 549
 QY 542 TGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGTGGGAAACCCGACAGACTATA 601
 DB 550 TGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGTGGGAAACCCGACAGACTATA 609
 QY 602 AAGATACAGCGGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTCTGTTCCG 653
 DB 610 AAGATACAGCGGTTTCCCGCTGGAAGCTCCCGCTGCGCTTTTCTGTTCCG 661

RESULT 12
 AV731085 735 bp mRNA linear EST 17-OCT-2000
 LOCUS AV731085 HTF Homo sapiens cDNA clone HTFAH04 5', mRNA sequence.
 DEFINITION AV731085
 ACCESSION AV731085
 VERSION AV731085.1 GI:10840506
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 735)
 AUTHORS Gu, Y., Peng, Y., Song, H., Huang, Q., Yang, Y., Gao, G., Xiao, H., Xu, X.,
 Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu,
 S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R.,
 Chen, J., Chen, Z., and Han, Z.
 TITLE Homo sapiens cDNA HTF clones
 JOURNAL Unpublished (2000)
 COMMENT Contact: Zeguang Han

Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919(ex.45)
 Fax: 86-21-50801922
 Email: hanzg@hgc.sh.cn
 This clone is available at CHGC in Shanghai.

FEATURES
 source 1..735
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HTFAH04"
 /clone_lib="HTF"
 /tissue_type="Hypothalamus"
 /dev_stage="Adult"
 /lab_host="SOLR"
 /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
 BASE COUNT 155 a 206 c 194 g 173 t 7 others
 ORIGIN

Query Match 86.4%; Score 637.6; DB 9; Length 735;
 Best Local Similarity 96.3%; Pred. No. 9.3e-174;
 Matches 673; Conservative 0; Mismatches 19; Indels 7; Gaps 2;

QY 40 GGGATCCCTAGAGTCGACCTCGAGCATGCAAGCTTGAGTATTCATAGTGTCACTAA 99
 DB 1 GGGATCCCTAGAGTCGACCTTCAGGCATGCAAGCTTGAGTATTCATAGTGTCACTAA 60
 QY 100 ATAGCTTGGCGTAATCATGGTTCATAGCTGTTTCTGTGTGAAATGTTATTCGGCTCACAA 159
 DB 61 ATAGCTTGGCGTAATCATGGTTCATAGCTGTTTCTGTGTGAAATGTTATTCGGCTCACAA 120
 QY 160 TTCCACACAACTACAGAGCGGGAAGCATAAAGTGTAAAGCCCTGGGTGCTAATAGTGA 219
 DB 121 TTCCACACAACTACAGAGCGGGAAGCATAAAGTGTAAAGCCCTGGGTGCTAATAGTGA 180
 QY 220 GCTAACTCACATTAATTCGGTTGCGCTCACTGCGCGCTTTCAGTGGGAAAGCTCTCGT 279
 DB 181 GCTAACTCACATTAATTCGGTTGCGCTCACTGCGCGCTTTCAGTGGGAAAGCTCTCGT 240
 QY 280 GCCAGCTGCATTAATGAATTCGGCCCAACGCGGGAAGAGCGGTTTTCGCTATTTGGCGCT 339
 DB 241 GCCAGCTGCATTAATGAATTCGGCCCAACGCGGGAAGAGCGGTTTTCGCTATTTGGCGCT 300
 QY 340 CTTCCGCTTTCCTCGCTCACTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGT 399
 DB 301 CTTCCGCTTTCCTCGCTCACTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGT 360
 QY 400 CAGCTCACTCAAGGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACGCGAGGAAGA 459
 DB 361 CAGCTCACTCAAGGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACGCGAGGAAGA 420
 QY 460 ACATGTGAGCAAAAGGCCAGCAAAAGGCCAGCAAAAGGCCAGCTTTCCTGGCGT 519
 DB 421 ACATGTGAGCAAAAGGCCAGCAAAAGGCCAGCAAAAGGCCAGCTTTCCTGGCGT 476
 QY 520 TTTTCCATAGGCTCGCGCCCTCGAGCATCAAAAATCGACGCTCAAGTCAGAGGT 579
 DB 477 TTTTCCATAGGCTCGCGCCCTCGAGCATCAAAAATCGACGCTCAAGTCAGAGGT 536
 QY 580 GCGGAAACCCGACAGACTATAAAGATACAGCGGTTTTCCTCGCTGGAAGCTCCCTCGTGC 639
 DB 537 GCGGAAACCCGACAGACTATAAAGATACAGCGGTTTTCCTCGCTGGAAGCTCCCTCGTGC 593
 QY 640 GCTCTCTCTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAA 699
 DB 594 GCTCTCTCTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAA 653
 QY 700 GGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCA 738
 DB 654 GGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCA 692

```
RESULT 13
AU081040 1039 bp mRNA linear EST 30-AUG-2001
LOCUS
DEFINITION
AU081040 Oncorhynchus mykiss Kidney infected by infectious
hematopoietic necrosis virus Oncorhynchus mykiss cdna clone KA9,
mRNA sequence.
ACCESSION
AU081040
VERSION
AU081040.1 GI:6431388
KEYWORDS
EST.
SOURCE
rainbow trout.
ORGANISM
Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 1039)
Kono,T., Sakai,M. and LaPatra,S.E.
Expressed Sequence Tag Analysis of Kidney and Gill Tissues from
Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious
Hematopoietic Necrosis Virus
Mar. Biotechnol. 2 (5), 493-498 (2001)
JOURNAL
COMMENT
Contact: Masahiro Sakai
Faculty of Agriculture
Miyazaki University
1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
Email: m.sakai@cc.miyazaki-u.ac.jp.
FEATURES
Location/Qualifiers
1..1039
/organism="Oncorhynchus mykiss"
/db_xref="taxon:8022"
/clone="KA9"
/clone_lib="Oncorhynchus mykiss Kidney infected by
infectious hematopoietic necrosis virus"
/tissue_type="Kidney infected by infectious hematopoietic
necrosis virus"
BASE COUNT 258 a 256 c 260 g 265 t
ORIGIN
Query Match 86.3%; Score 637; DB 9; Length 1039;
Best Local Similarity 100.0%; Pred. No. 1.6e-173;
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 102 AGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAATTTGTTATCCGCTCACAAAT 161
DB 345 AGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAATTTGTTATCCGCTCACAAAT 404
QY 162 CCACACAACATACAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCTTAATGAGTGAGC 221
DB 405 CCACACAACATACAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCTTAATGAGTGAGC 464
QY 222 TAACACATTAATTTGGGTGGCGTCACTGCCCGCTTTCAGTCGGGAACCTGTCGTGC 281
DB 465 TAACACATTAATTTGGGTGGCGTCACTGCCCGCTTTCAGTCGGGAACCTGTCGTGC 524
QY 282 CAGCTGCATTAATGAATCGGCCCAACGCGGGAGAGCGGTTTTCGCTATTGGCGGCTCT 341
DB 525 CAGCTGCATTAATGAATCGGCCCAACGCGGGAGAGCGGTTTTCGCTATTGGCGGCTCT 584
QY 342 TCGGCTTCCTCGCTCACTGACTCGCTGGGTTCGCTGCTGGGTTCGCGTTCGCGGAGCGGTAFCA 401
DB 585 TCGGCTTCCTCGCTCACTGACTCGCTGGGTTCGCTGCTGGGTTCGCGTTCGCGGAGCGGTATCA 644
QY 402 GCTCACTCAAGGGGGTAATACGGTTATCCACAGAATCAGGGGATACGCGAAGAAAGAAC 461
DB 645 GCTCACTCAAGGGGGTAATACGGTTATCCACAGAATCAGGGGATACGCGAAGAAAGAAC 704
QY 462 ATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCGGGTTCGTCGCGCTTT 521
DB 705 ATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCGGGTTCGTCGCGCTTT 764
QY 522 TTCCATAGGCTCCGCCCTTGACGAGCATACACAAAATCGACCTCAAGTCAGAGGTGG 581
```

```
DB 765 TTCCATAGGCTCCGCCCTGACGAGCATCAAAAAATCGAGCTCAAGTCAGAGGTGG 824
QY 582 CGAAACCCGACAGGACTATAAAGATACCAAGCGTTTCCCTCGGAAGCTCCCTCGTGCGC 641
DB 825 CGAAACCCGACAGGACTATAAAGATACCAAGCGTTTCCCTCGGAAGCTCCCTCGTGCGC 884
QY 642 TCTCTGTTCGGACGCTGCCGCTTACCGGATACCTGTCCGCCCTTTCTCCCTTCGGGAAGC 701
DB 885 TCTCTGTTCGGACGCTGCCGCTTACCGGATACCTGTCCGCCCTTTCTCCCTTCGGGAAGC 944
QY 702 GTGGCGCTTTCATAGCTACGCTGCTAGGTATCTCA 738
DB 945 GTGGCGCTTTCATAGCTACGCTGCTAGGTATCTCA 981
RESULT 14
AU081124 1089 bp mRNA linear EST 30-AUG-2001
LOCUS
DEFINITION
AU081124 Oncorhynchus mykiss Kidney infected by infectious
hematopoietic necrosis virus Oncorhynchus mykiss cdna clone KG'12,
mRNA sequence.
ACCESSION
AU081124
VERSION
AU081124.1 GI:6431472
KEYWORDS
EST.
SOURCE
rainbow trout.
ORGANISM
Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 1089)
Kono,T., Sakai,M. and LaPatra,S.E.
Expressed Sequence Tag Analysis of Kidney and Gill Tissues from
Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious
Hematopoietic Necrosis Virus
Mar. Biotechnol. 2 (5), 493-498 (2001)
JOURNAL
COMMENT
Contact: Masahiro Sakai
Faculty of Agriculture
Miyazaki University
1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
Email: m.sakai@cc.miyazaki-u.ac.jp.
FEATURES
Location/Qualifiers
1..1089
/organism="Oncorhynchus mykiss"
/db_xref="taxon:8022"
/clone="KG'12"
/clone_lib="Oncorhynchus mykiss Kidney infected by
infectious hematopoietic necrosis virus"
/tissue_type="Kidney infected by infectious hematopoietic
necrosis virus"
BASE COUNT 245 a 297 c 295 g 251 t
ORIGIN
Query Match 86.3%; Score 637; DB 9; Length 1089;
Best Local Similarity 100.0%; Pred. No. 1.6e-173;
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 102 AGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAATTTGTTATCCGCTCACAAAT 161
DB 81 AGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAATTTGTTATCCGCTCACAAAT 140
QY 162 CCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCTTAATGAGTGAGC 221
DB 141 CCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCTTAATGAGTGAGC 200
QY 222 TAACACATTAATTTGGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 281
DB 201 TAACACATTAATTTGGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 260
QY 282 CAGCTGCATTAATGAATCGGCCCAACGCGGGAGAGCGGTTTTCGCTATTGGCGGCTCT 341
DB 261 CAGCTGCATTAATGAATCGGCCCAACGCGGGAGAGCGGTTTTCGCTATTGGCGGCTCT 320
```

[illegible]

Search completed: August 1, 2002, 08:39:59
Job time: 6490 sec

Query Match 86.3%; Score 637; DB 9; Length 1163;
Best Local Similarity 100.0%; Pred. No. 1.7e-173;
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 102 AGCTTGGCGTAATCATGGTCATAGCTCTTTCCCTGTGGAATTTCTTTCGCTCACAATT 161

us-10-014-743-3.rst

Thu Aug 1 12:08:51 2002